

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:26 ; Search time 34.7285 Seconds
(without alignments)
11051.549 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532
Perfect score: 3615
Sequence: 1 atgagccatgagatgatgttg.....ttggaatcatctgaggtcttc 1995

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=x1p
-Q=/cgn2.1/USPTO.spool.p/US10029345/runatc.21062004.122816.4205/app.query.fasta_1.3278
-DB=pir.78 -OPMT=fastan -SUFFIX=rxpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdd -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORHEX= -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US10029345 -QCGN_1_1_78 @runatc.21062004.122816.4205 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESUBSTR -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: pir.78:.*
2: pir1:.*
3: pir2:.*
4: pir3:.*
5: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	515	14.2	619	2	T15969
2	443	12.3	394	2	A56115
3	433	12.0	367	1	S24411
4	423	11.7	367	2	S52285
5	420	11.6	367	1	S29090
6	418	11.3	314	1	A57126
7	409	11.3	314	2	B57126
8	393.5	10.9	393	2	A56947
9	360.5	10.0	384	2	T13880
10	345	9.5	303	2	T16405
11	294.5	8.1	365	2	T132494
12	239.5	6.6	272	2	T18915
13	233.5	6.5	226	2	T13180
14	225.5	6.2	186	2	T16056

15	223.5	6.2	330	2	T139698	protein tyrosine p
16	218.5	6.0	771	2	T47666	phosphatase-like p
17	208.5	5.8	220	2	T37885	low-molecular-mass
18	205.5	5.7	866	2	P88481	protein C16A3.1 [1
19	205	5.6	223	2	T19317	protein tyrosine p
20	204	5.6	278	2	T39517	dual-specificity p
21	202	5.6	142	2	T03074	dual specificity p
22	202	5.6	364	1	S31304	protein-tyrosine-p
23	196	5.4	185	1	A47196	dual specificity p
24	195	5.4	223	2	G84458	probable protein p
25	193	5.3	205	2	T49364	protein tyrosine p
26	189	5.2	489	1	S58725	dual specificity p
27	185	5.1	807	1	S44538	probable protein-c
28	183.5	5.1	276	2	T48906	protein-tyrosine-p
29	178	4.9	204	2	T17802	hypothetical prote
30	177	4.9	580	2	T18439	hypothetical prote
31	172	4.8	600	2	T18446	hypothetical prote
32	170.5	4.7	292	2	S41012	hypothetical prote
33	160	4.4	1132	2	T49403	related to protein
34	157	4.3	209	1	S48459	probable dual spec
35	154.5	4.3	5262	2	T03454	ALR protein - huma
36	154	4.3	1189	2	S56852	hypothetical prote
37	151.5	4.2	597	1	S43743	probable dual spec
38	151	4.2	1141	2	A44093	CGMP-inhibited CAM
39	150.5	4.2	169	2	T30684	probable dual spec
40	148.5	4.1	574	2	T43556	Wiskott-Aldrich sy
41	147	4.1	1168	2	S49915	extensin-like prot
42	146.5	4.1	272	2	T19418	hypothetical prote
43	145.5	4.0	171	1	T36845	dual specificity p
44	145.5	4.0	171	2	T28522	probable dual spec
45	145.5	4.0	171	2	B72161	J1L protein - vari

ALIGNMENTS

RESULT 1

T15969

hypothetical protein F08B1.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T15969

R/Chisoe, S.

submitted to the EMBL Data Library, July 1995

A/Description: The sequence of C. elegans cosmid F08B1.

A/Reference number: Z18439

A/Accession: T15969

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-619 <CHI>

A/Cross-references: EMBL:U23178; NID:G726421; PID:G726422; PIDN:AC46719.1; CESP:F08B1.1

A/Experimental source: strain Bristol N2

A/Gene: CESP:F08B1.1

A/Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Alignment Scores:

Pred. No.:	7.36e-24	Length:	619
Score:	515.00	Matches:	165
Percent Similarity:	45.32%	Conservative:	72
Best Local Similarity:	31.55%	Mismatches:	173
Query Match:	14.25%	Indels:	113
DB:	2	Gaps:	19

US-10-029-345A-108_COPY_538_2532 (1-1995) x T15969 (1-619)

QY	367	GGTGGTTTCTGAGTTCCTCTGTTTCCTGCTCTGTGAAGAAATCC-----	420
DB	99	GL[GL]PheLysInhheAaIngnInrPrGlnLeuCySeLserGInuGlyMet 118	
QY	421	ACTTGAATCCCTACCTGCATTTCTGAGCTTCTTA-----CCTGTGGCAACATTGGG 474	
DB	119	ThArgLeuProGlnSerLeuSerGInProCyLeuSerGInProThrArgLysP---Gly 137	

```

QY 475 CCAACCCGAATTCCTTCATCTTATCTGGCTCCGACGAGATGCTCTCAACAGAG 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 lIethleuileIethProsmilleTyrluenglySerGlnIleasPserleuaspGluThr 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 535 CTGATACAGACAAATGGATGGATGTTATGTTAAATGACAGCTATACCTGCGAAAGCT 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 MetLeuaspAlaLeuaspIleSerValIleAsnLeuSerMetThrCysProlySer 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 595 GACTTATATCCCGAG--TCTCATTTCCGCGTGGCTGTGAATGACAGCTTTTGTAG 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 ValCysIleIysGlnaspIlyasnPhenethArgIleProValaAsnSerTyrgInglu 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 AAAATTTGCGCGTTGACAAATCATATGATTTCATTTGAAAGAAAGAAAGCTCCAAAT 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 LysLeuSerProlyrPheProMetAlaTyrgIuIheugIuIyScyAsnArgAlaGly 217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 712 GGATGTGTTAGTGCATCTTTTATGCTGGATCTCCGCTCCGACCATGCTATGCGC 771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 LysIyScyLeuIleIhScyLeuAlaGlyIleSerArgSerProThrIleAlaIleSer 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 772 TACATCATGAAGAGATGACATGTCTTTAGATGATGACATTTGTGAAAGAAA 831
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 TyrlIleMetArgTyrlMetIySerMetGlySerAsnaspAlaTyrgTyrgValIyScyIuArg 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 832 AGACCTATATATCTCCAACTTCATTTTGTGGCCCACTCTGACATATGAGAAG 891
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 ArgProSerIleSerProAsnPhenaspMetGlyGlnLeuIuIyTyrgIuIuVal 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 892 ATTAAGAACAGACTGAGATCAGGCGCAAAAGCAATCAAGTCTGACCTGGAG 951
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 Leu-----IleIyAsnPhisValIleuaspTyrgAsn 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 952 AAGCAATATACCTGCTCCCTGTCTGTCTGACAGGCTGACAGAAAAGCAGAGCGCCCTC 1011
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 GlnAlaSerArgProIhAsnArgIhMetAspTyrgTyrgIyProSerAsnIySerPro 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1012 AGTCCACCTGTGCGACCTGTCTGATCTGACAGGCGACGAGCAAGAAAGCCGCTGATCCC 1071
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 ---ProlyValrProlySerSerAlaSerSerAsnCysValrPheProIySerThrIhAsp 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1072 GCCAGCGTCCCGACGCTGCGCCAGCGCTGCGCTGTGTAGAGACAGCCCGCTGTGTA 1111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 GluSerSerProSerSerProSerValSerGluIySerAlaIleSerGluPro----- 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1132 CAGGCGCTAGTGGCGTGCACCTGTCCGACAGCGCTGGAAGACAGCAATTAAGCTCAAG 1191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 343 ----- 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1192 CGTCTCTCTCTGTGATATCAATCAATTCATATTCAGCGACGATGGACATCTCTTA 1251
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 344 -----GluThrSerSerSerAlaIleSer--- 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1252 CATGGCTTCTCTCATCAGAGATCTTTGGAATTCACAACTTCACATCTGTGAT 1311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 -----SerSerSerThrAlaSerAlaProProSerMetProSerThrSerGluGln 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1312 GGGACCAACAGACTATGCGACATCTCCCTGTTCAGGAACATAGGAGCAGACTCCGAA 1371
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 GlyThrSer-----SerGlyThrValAsnValAsnGlyIyAsnArgAsnMet 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1372 ACCAGT-----CTGATAGAGAGAGAGCC--AGCATCCCAAGAGATGACG 1416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 ThrMetAsnLeuGlyLeuProIhAsnArgProlyAlaIleuGlyLeuProSerArgIleGly 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1417 ACCGCT-----AGGCTTCAGACAGC--CAGAGCAAGGATTTGCTGCTACAA 1464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 404 ThrSerValaIleGluLeuProSerProSerThrGluLeuSerArgIleuSerPhe--AsnG1 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1465 ACCGACGACGAGTGGACCGCCCGACAGAGTCCCTTTATCTTCACATGATGAGAGTGGAGC 1524
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 yProGluAlaIleAlaProSerThr---ProIleLeu----- 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1525 GTGAGAGACAATTAACACACACAGCTTCCTTTTGGCGCTTCCACGACGACGACCTTC 1584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 435 -----AsnPhenAsnProCysPheAsnSerProIleIleProValaIleSerSerSe 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1585 ACGAAGCTGCTGCTGAGCGCTTAAGAGGCTGGACACTGATATCTTGGCCCCCAGACC 1644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 rArgIuAlaIle----- 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1645 TCTACCCCTCCCTGACAGCAGCTGATTTTCCACAGAGTCTTCACTTACTCT 1704
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 457 -LeuThrIeuProThrProAlaIa-----SerSerSerSerThrSe 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1705 GCGTCAGCCA-----TTCAGGAGGACAGTCCAGCT 1734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 471 rSerGluProSerPheAsnPhenSerSerPheGluSerSerSerSerIleValVa 491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1735 TACTGTGCTACACTGACGACGCTGCCACTTTCGCGAGACCAAGTCTATTCTGTGCGC 1794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 491 lGluAsnProPhePheAlaSerThrGluValrProAlaGlySerSerSerIleSerThrPr 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1795 AGCGGAGAAAGCCAGTACACAGAGCTGACTCGCGGCGAGCTGCGCATGAAGAGACCC 1854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 511 oSerGly---SerGluSerThrProAlaSerAlaSerSerSerAlaIleSerArgCysAr 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1855 TTTGAAA 1861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 530 gMetIys 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 2

```

A56115 dual specificity phosphoprotein phosphatase (RC 3.1.3.-) 4 - human
N:Alternate names: dual specificity phosphatase HVH2
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Apr-1998
R:Accession: A56115
R:Guan, K.L.; Butch, E.
J. Biol. Chem. 270, 7197-7203, 1995
A:Title: Isolation and characterization of a novel dual specific phosphatase, HVH2, whic
A:Reference number: A56115; PMID:95221370; PMID:7535768
A:Accession: A56115
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-394 <GUA>
A:Cross-references: GB:U21108
C:Genetics:
A:Gene: GDB:DUSP4; HVH2; MKP-2
A:Cross-references: GDB:433893
A:Map position: 8p21-8p11.2
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHN-type dual specificity
C:Keywords: phosphoprotein; phosphoric monoester hydrolase
F:203-334/Domain: VHN-type dual specificity phosphoprotein phosphatase homology <VHN>
F:280/Active site: Cys (phosphocysteine intermediate) #status predicted
F:286/Binding site: substrate phosphate (Arg) #status predicted

```

Alignment Scores:

```

Pred. No.: 1,75e-19 Length: 394
Score: 443.00 Matches: 120
Percent Similarity: 50.79% Conservative: 74
Best Local Similarity: 31.41% Mismatches: 142
Query Match: 12.25% Indels: 46
DB: 2 Gaps: 14

```

US-10-029-345A-108_COPY_538_2532 (1-1995) x A56115 (1-394)

```

QY 55 CTGGAAGTGAACGGAAGAAAGTGTCTTAATGATAGCGCGCATTTTGTGAAATCAAT 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39 LeuProSerGly---GlyIyScyLeuIleuAsnAspCysArgProPheIleuAlaIleSer 57
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 115 ACATCCACATTTTGAAGCCATTAATCAACTGCTCCCAACTTATGAAGCAAGTTG 174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 AlaglyTyrlleuIySerValaAsnValArgCysAsnThrIleValArgArgArg--- 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 CAACAGACAAAGTGAATTAATTCAGAGCTCATC-----CAGCATTCAGCGAAACATTAAG 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db	77	AlaIys6GlySerValSerLeuGlnGlnIleLeuProAlaGluGluValAlaArg	96
QY	229	GTTCACATGATGATGACAGACGAGGTGTGATTACGATCAAAAGCTCCCAAGATGTTGCC	288
Db	97	LeuArgSerGlyLeuGlyTyrSerAlaValAlaIleValIleTyrArgGlyArgSerProAlaGlu	116
QY	289	TCTCTCTCTTACAGACTGTTTTCTCAGCTACTCTCTGGGTAACTGGAGAAAGCTTC---	345
Db	117	SerLeuArgGluAerSerThrValSerLeuValAlaGlnAlaLeuAaGArgValaGlu	136
QY	346	---AAGCTGTTCACCTGTCGACAGGTGGGTGGTGGTGGTCTGCTGGTGGTGGTGGC	402
Db	137	ArgThrAerPrlIecy6LeuLeuLysGlyGlyTyrGlnAArgPheSerSerGlyTyrProGlu	156
QY	403	CTCTGTGAGGAAAAATCCACTCTA-----GTCCCTACTGTCGATTTCT	444
Db	157	PheCySerLys7ThrLysAlaLeuAlaIleProProProValProProSerAlaThr	176
QY	445	CAGCCT-----TGC---TTACCTGTGTGCCAATC---GGGCA	477
Db	177	GluProLeuAerLeuGlyCySerSerSerGlyThrProLysHisAerGlnGlyGlyPro	196
QY	478	ACCGCAATTTCTCCCAACTTTTATCTTGGTGGCCGACAGATATGCTCCACAAAGAGTGG	537
Db	197	ValGlnIleLeuProPheLeuTyrLeuGlySerAlaTyrHisAlaAlaArgAerPheC	216
QY	538	ATACAGCAGAAATGGGATGTTATGTGTATTAATGACAGCTATACCTGTCCAAAGCTGAC	597
Db	217	LeuAerAlaLeuGlyIleThrAlaLeuLeuAerHisSerAerProAerHis	235
QY	598	TTTATCCCCGAGCTCATTTCTCTGCGTGGCTGCTGGAATGACAGCTTTTGGAGAAATTT	657
Db	236	PheGlnGlyHisTyrGlnTyrLysCysLysProValGluAerHisValAerPrlLe	255
QY	658	TTGCGCGGTGTGACAAATCAGTATGATTCATTCATGAGAAAGAAAGCCTCCAAAGAGT	717
Db	256	SerSerThrPheMetGlnAlaIleGlnTyrIleAerAlaValLysAerCyAerGlyArg	275
QY	718	GTTCTAGTGCACCTGTTTATGCTGGAGATCTCCGCGCTCCGACACATCGCATGCTACATC	777
Db	276	ValLeuValHisCy6GlnAlaGlyIleSerThrAerAlaThrIleCysLeuAlaTyrLeu	295
QY	778	ATGAAGAGATGACATGCTTTTACATGAAGCTTACAGATTTGTGAAGAAAGAAAGACCT	837
Db	296	MetCysLysAerGlyValArgLeuGlnGluAlaPheGlnPheValLysGlnAerGlySer	315
QY	838	ACTATATCTCCAACTTCAATTTCTGGGCCCACTCCGACATATGAGAAAGATTAAG	897
Db	316	IleIleSerProHisPheSerPheMetGlyGlnLeuLeuGlnPheGlnSerGlnAlaLeu	335
QY	898	AACAGAGATGAGCATCAGAGCCCAAGAGCAAAAGCTCAAGCTGTGTCACCTGTGAGAGCA	957
Db	336	AlaThrSerCy6AlaIleGluAlaAer-----Pro	346
QY	958	AATGAACCTGTCTCCCTGTCTGTCTCAGAGGCTGACAGAAAAGCGACCCCTCAGTCCA	1017
Db	347	SerGlyProLeu-----ArgGlnAerGlyLysThrProAlaThrPro	360
QY	1018	CCCTGTGCCAGCTGCTACTCACTCAGAGGACAGACAGAAAGCCCGTGCATCCGCGCAGC	1077
Db	361	-----ThrSerIlePheValPheSerPheProVal---SerValGly	373
QY	1078	GTGCCAGACGTGCCAGCGGTGCGCTGTGTTAGAGACAGCGCCGTGATCAGAGCG	1137
Db	374	ValHisSerAlaProSerSerLeuProGlyLeu-----HisSerProIleThrThrSer	391
QY	1138	CTCAGT 1143	
Db	392	ProSer 393	

[illegible]

Db 171 GlnGlyGlyProValGluIleLeuSerPheLeuTyrLeuGlySerAlaTyrHisAlaSer 190
 QY 526 AACAAAGAGCTGATACGACGAATGGATTGGTTATGTTAAATCCGATATACCTGT 585
 Db 191 ArgLysAspMetLeuAspAlaLeuGlyIleThrAlaLeuIleAsnValSerIleAsnLys 210
 QY 586 CCAAGAGCTGACTTATATCCCGAGTCTCATTTCTCCGTCGTCCTGTGAATGACGCTTT 645
 Db 211 ProAsn---HisPheGluGlyHisTyrGlnTyrLysSerIleProValGluAspAsnHis 229
 QY 646 TGTGAGAAATTTTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 705
 Db 230 LysAlaAspLysSerSerTyrPheAsnGlnAlaIleAspPheIleAspSerIleLysAsp 249
 QY 706 TCCAAATGATGTGTCTAGTGCATGTTAGTGGATCTCCGTCGTCGTCGTCGTCGTCG 765
 Db 250 AlaGlyGlyArgValPheValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCys 269
 QY 766 ATGCGCTACATCATGACGATGACGATGACATGTCCTTATGATGAGCTTACAGTTGTGAA 825
 Db 270 LeuAlaTyrLeuMetArgThrAsnArgValLysLeuAspGluAlaPheGluPheValLys 289
 QY 826 GAAAAAGAGCTGATATCTCCAAACTGTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 885
 Db 290 GlnArgArgSerIleIleSerProAsnPheSerPheMetGlyGlnLeuGlnPheGln 309
 QY 886 AAGAAGATTAAAGACCACTGACGATCAGAGGCGCAAGAGCAAACTCAAGTGTGTCAC 945
 Db 310 SerGlnVal-----LeuAlaProHisCysSer----- 318
 QY 946 CTGAGAGAGCCAAATGAACTGTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1005
 Db 319 ---AlaGlnAlaGlySerProAlaMetAlaValLeuAspArgGlyThrSerThrThrThr 337
 QY 1006 CCCCTCAGTCCACCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1065
 Db 338 ValPheAsnPheProValSer-----IleProVal 347
 QY 1066 CATCCCGCCAGC 1077
 Db 348 HisProThrAsn 351

RESULT 4

S52265
 dual specificity phosphatase (EC 3.1.3.-) 1 - rat
 N/Alternate names: protein-tyrosine-phosphatase CL100; protein-tyrosine-phosphatase, not
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999
 R/Mda, M.; Schlegel, W.; Arkinsteil, S.
 submitted to the EMBL Data Library, January 1995
 A/Description: Pathways regulating CL100 gene expression in pituitary cells.
 A/Reference number: S52265
 A/Accession: S52265
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-367 <MNU>
 A/Cross-references: EMBL:X84004, NID:G642264, PID:CAA58828.1, PID:G642265
 C/Suprafamily: dual specificity phosphoprotein phosphatase 1; WH1-type dual specificity
 C/Keywords: phosphoprotein; phosphoric monoester hydrolase
 F/181-312/Domain: WH1-type dual specificity phosphoprotein phosphatase homology <WH1>
 F/258/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/264/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:

Pred. No.: 2.88e-18 Length: 367
 Score: 423.00 Matches: 115
 Percent Similarity: 46.05% Conservative: 54
 Best Local Similarity: 31.34% Mismatches: 148
 Query Match: 11.70% Indels: 50
 DB: 2 Gaps: 9

US-10-029-345a-108_copy_538_2532 (1-1995) x S52265 (1-367)

QY 43 TTGGTGCTCTGTGAAAGTGAACGAAAGAGCTGTAATTGATACCGGCACTTT 102
 Db 13 LeuArgAlaLeuLeuArgGluLysArgAlaIleGlnCysLeuLeuAspCysArgSerPhe 32
 QY 103 GTGGAATACATACATCCCATTTTGGAGCCATTATATCACTGCTCAAGCTTATG 162
 Db 33 PheAlaPheAsnAlaGlyHisIleValGlySerValAsnValArgPheSerThrIleVal 52
 QY 163 AAGCAAGCTTGAAC-----CAGACAAAGCTGTTATTCAGAGCTCATC 207
 Db 53 ArgArgArgAlaLysGlyAlaMetGlyLeuGlnHisIleValProAsnThrGlnLeu--- 71
 QY 208 CAGCATTCAGGAAACATTAAGCTTGATGATGATGACGACGAGGCTGATGATGATGAT 267
 Db 72 -----ArgGlyArgLeuLeuAlaGlyAlaTyrHisAlaValValLeuLeuAsp 87
 QY 268 CAAGAGCTCCCAAGATGTGCT 327
 Db 88 GluArgSerAlaAlaLeuAspGlyAlaLysArgAspGlyThrLeuAlaLeuAlaIleGly 107
 QY 328 AAAGCTG-----GAGAAAGCTTCAACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 381
 Db 108 AlaLeuCysArgGluAlaArgSerThrGlnValPhePheLeuGlnGlyTyrGluAla 127
 QY 382 TTCCTGCTGTTTCCCTGCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 417
 Db 128 PheSerAlaSerCysSerProGluLeuCysSerLysGlnSerThrProMetGlyLeuSerLeu 147
 QY 418 -----TTCACCTCTAGTCCCTTACCTGACATTTCTCAGCCTTCACTGTCGTCGTCGTC 471
 Db 148 ProLeuSerThrSerValProAspSerAlaGlnSerGlyCysSerSerCysSerThrPro 167
 QY 472 -----GGGCAACCCGAATCTCTCCCAATCTTTATCTTGGTCGTCGTCGTCGTCGTCG 516
 Db 168 LeuTyrAspGlnGlyGlyProValGluIleLeuSerPheLeuTyrLeuGlySerAlaTyr 187
 QY 517 GATGTCCTCAACAGACGCTGATACAGCAGATGGAATGGTATGTTATGTTAAATGTCAGC 576
 Db 188 HisAlaSerArgLysAspMetLeuAspAlaLeuGlyIleThrAlaLeuIleAsnValSer 207
 QY 577 TATACCTGTCCAAGAGCTGATATCCCGAGTCTATTTCCGTCGTCGTCGTCGTCGTCGTCG 636
 Db 208 AlaAsnCysProAsn---HisPheGlnGlyHisTyrGlnTyrLysSerIleProValGlu 226
 QY 637 GACAGCTTTTGTGAAATTTTGCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 696
 Db 227 AspAsnHisLysAlaAspLysSerSerTyrPheAsnGlnAlaIleAspPheIleAspSer 246
 QY 697 GCAAAAGCTTCATGATGATGTGTTAGTCACTGTTAGCTGGGATCTCCGTCGTCGTCGTCG 756
 Db 247 IleLysAspAlaGlyArgValPheValHisCysGlnAlaGlyIleSerArgSerAla 266
 QY 757 ACCATGCTTATCCGCTCATCATGATGAAAGATGACATGCTTTAGATGATGATGATGATGAT 816
 Db 267 ThrIleCysLeuAlaTyrLeuMetArgThrAsnArgValLysLeuAspGluAlaPheGlu 286
 QY 817 TTGTGGAAGAAAGAGCTTATCTATCTCAAACTTCAATTTTGGCCCAATCTCTG 876
 Db 287 PheValLysLeuAlaArgArgSerIleIleSerProAsnPheSerPheMetGlyGlnLeuLeu 306
 QY 877 GACTATGAGAAAGATTAAAGACCACTGAGCATCAGGCGCAAGAGCAAACTCAAG 936
 Db 307 GlnPheGlnSerGlnVal-----LeuAlaProHisCysSer----- 318
 QY 937 CTGCTGCACTGAGAGCAAAATGAACTGTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 996
 Db 319 -----AlaGluAlaGlySerProAlaMetAlaValLeuAspArgGlyThrSer 334
 QY 997 AGCGAGAGCCCTTCACTGACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1056
 Db 335 ThrThrThrValPheAsnPheProValSer----- 344

QY 1057 AGGCCGTCATCCCGCCAGC 1077
 |||||
 Db 345 lIeProValHisProThraen 351

RESULT 5

S23090

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human

N/Alternate names: protein-tyrosine-phosphatase CIL100; protein-tyrosine-phosphatase, not

C/Species: Homo sapiens (man)

C/Date: 25-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999

C/Accession: S23090; #sequence_revision 02-May-1994 #text_change 11-Jun-1999

R/Keyes, S.M.; Emalle, E.A.

Nature 359, 644-647, 1992

A/Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine

A/Reference number: S23090; PMID:93024952; PMID:11406996

A/Accession: S23090

A/Molecule type: mRNA

A/Residues: 1-367 <KEY>

A/Cross-references: EMBL:X68277; NID:g29980; PIDN:CAA48338.1; PID:g29981

R/Keyes, S.P.; Hakes, D.J.; Martelli, K.J.; Dixon, J.E.

J. Biol. Chem. 269, 3596-3604, 1994

A/Title: Isolation and characterization of a human dual specificity protein-tyrosine pho

A/Reference number: A53052; PMID:94148864; PMID:8106404

A/Accession: A53052

A/Molecule type: DNA

A/Residues: 1-367 <KEY>

A/Experimental source: Leukocyte

A/Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804,

C/Genetic:

A/Genes: GDB:DUSP1; PTPN10

A/Cross-references: GDB:136197; OMIM:600714

A/Map position: 5q34-5q34

C/Superfamily: dual specificity phosphoprotein phosphatase 1, vhl-type dual specificity

C/Keywords: heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced p

F/181-312/Domain: vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>

F/258/Active site: Cys (phosphocysteine intermediate) #status predicted

F/264/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:

Pred. No.:	4.38e-18	Length:	367
Score:	420.00	Matches:	113
Percent Similarity:	46.69%	Conservative:	56
Best Local Similarity:	31.22%	Mismatches:	159
Query Match:	11.62%	Indels:	34
DB:	1	Gaps:	7

US-10-029-345a-108_copy_538_2532 (1-1995) x S23090 (1-367)

QY 43 TTGGTGCTCTCTGGAAGTGAACGAAAGTGTGTAATGATACCGCCATT 102
 |||||
 Db 13 LeuArgAlaLeuLeuGlyGluArgAlaAlaGlnCysLeuLeuSerPhe 32

QY 103 GTGGAATACATATCATCCACATTTTGGAGCCATTAATATCACTGCTCAAGCTTATG 162
 ::|||
 Db 33 PheAlaPheHisAlaGlyHisIleAlaGlySerValHisAlaArgPheSerThrIleVal 52

QY 163 AAGCGAAGTTGGCAACAGACAAAGTTTAATTAACAGCTCATCCAGCTTACGCGAA 222
 ::|||
 Db 53 ArgArgArgAlaValGlyAlaMetGlyLeuGlnHisIleValProAlaAlaGluLeuArg 72

QY 223 CATTAAGTTGACATGATGTTGACATGACGATGATGATTAATTAACGATCAAGCTCCCAAGAT 282
 ::|||
 Db 73 GlyArgLeuLeuAlaGlyAlaValHisAlaValAlaLeuLeuSerPheGluArgSerAlaAla 92

QY 283 GTTGCCCTCTCTCTTCACATGTTTTCACATGTTTTCACATGTTTTCACATGTTTTC 336
 ::|||
 Db 93 LeuArgPheAlaValArgSerPheGlyThrLeuAlaLeuAlaAlaGlyAlaValCysArgGlu 112

QY 337 AAGAGCTTCACTGTTTTCACATGTTTTCACATGTTTTCACATGTTTTCACATGTTTTC 396
 ::|||
 Db 113 AlaArgAlaAlaGlnValPhePheLeuLeuGlyGlyGlyGluAlaPheSerAlaSerCys 132

QY 397 CCGGCGCTCTGTGAAGAAAA-----TCAACTCTA 426

Db 133 ProGluLeuCysSerIysGlnSerThrPrometGlyLeuSerLeuProLeuSerThrSer 152
 |||||
 QY 427 GTCCCTACCTGACATTTCTACAGCTTGCTTACCTGTTCCCAACTT----- 471
 |||||

Db 153 ValProAspSerAlaGluSerGlyCysSerSerCysSerThrProLeuThrArgGlnGly 172
 |||||

QY 472 GGGCGAACCGAATGTTTCCCAATCTTTATCTTGGTCCGACGAGATGTCCTCAACAG 531
 |||||

Db 173 GlyProValGluLeuProPheLeuTyrlleGlySerAlaThrHisAlaSerArgGly 192
 |||||

QY 532 GACGATACACAGACGATGATGTTGTTATGTTTAAATGCGACTATACCTGTCACAG 591
 ::|||

Db 193 AspMetLeuAlaPheAlaLeuGlyIleThrAlaLeuHisAlaValSerAlaCysProAsn 212
 ::|||

QY 592 CCGTACTTATCCCGAGTCTCATTTCCGCGTGGCTGTGATGACAGCTTTGTGAG 651
 ::|||

Db 213 ---HisPheGluGlyHisIleTyrlleGlySerIleProValGluAlaPheHisIleVal 231
 ::|||

QY 652 AAAATTTTCCCGTGGTGGACAAATGATGATTTGATGAGAAAGCAAGCTCCAAAT 711
 ::|||

Db 232 AspIleSerSerThrPheAsnGlnAlaIleAspPheIleAspSerIleTyrlleAsnAlaGly 251
 ::|||

QY 712 GGATGTTTCTAGTACATGTTTATGTTGATGATGATGATGATGATGATGATGATGATG 771
 ::|||

Db 252 GlyArgValPheValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCysLeuAla 271
 ::|||

QY 772 TACATCATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 831
 ::|||

Db 272 TyrLeuMetArgThrAsnArgValIleValAspGluAlaPheGluPheValIleGlnArg 291
 ::|||

QY 832 AGACCTACATATATCTCCAACTTCAATTTTGGGCACTCCTGACATATGAGAAAG 891
 ::|||

Db 292 ArgSerIleIleSerProAsnPheSerPheMetGlyGlnLeuLeuGlnPheGluSerGln 311
 ::|||

QY 892 ATTAAGAACACAGACTGACATGACGAGGCGCAAAAGCAAACTCAAGCTGTGACCTGAG 951
 ::|||

Db 312 Val-----LeuAlaProHisCysSer-----Ala 319
 ::|||

QY 952 AAGCGAATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
 ::|||

Db 320 GluAlaGlySerProAlaMetAlaValLeuAspArgGlyThrSerThrThrValPhe 339
 ::|||

QY 1006 -----CCCTCAGTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1059
 ::|||

Db 340 AsnPheProValSerIleProValHisSerThrAsnSerAlaLeuSerTyrlleGlnSer 359
 ::|||

QY 1060 CCCGTG 1065
 |||||

Db 360 ProIle 361
 |||||

RESULT 6

A57126

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 2 - human

N/Alternate names: mitogen-induced nuclear tyrosine phosphatase; phosphatase of activator

C/Species: Homo sapiens (man)

C/Date: 03-Nov-1995 #sequence_revision 02-Jul-1996 #text_change 11-Jun-1999

C/Accession: A57126

R/Rohan, P.J.; Davies, P.; Moskaluk, C.A.; Kearns, M.; Kruttsch, H.; Siebenlist, U.; Kelly,

Science 259, 1763-1766, 1993

A/Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.

A/Reference number: A57126; PMID:93206122; PMID:7681221

A/Accession: A57126

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-314 <ROH>

A/Cross-references: GB:L11329; NID:g559539; PIDN:AAA50779.1; PID:g292376

A/Genes: GDB:DUSP2

A/Cross-references: GDB:139200

A/Map position: 2q11-2q11

A/Function: catalyzes the hydrolysis of peptidyl-phosphoserine, -phosphothreonine, and

C:Superfamily dual specificity phosphoprotein phosphatase 1, VH1-type dual specificity
 C:Keywords: nucleus, phosphoprotein, phosphoric monoester hydrolase
 E:140-160/Region: nuclear location signal
 F:180-311/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
 F:257/Active site: Cys (phosphoserine intermediate) #status predicted
 F:263/Binding site: substrate phosphatase (Arg) #status predicted

Alignment Scores:
Prod vs

Seed: 1011	5.75e-18	Length:	314
Score:	418.00	Matches:	109
Percent Similarity:	50.83%	Conservative:	45
Best Local Similarity:	35.97%	Mismatches:	105
Query Match:	11.56%	Indels:	44
DB:	1	Gaps:	0

02-106-315A-108_COPY_538_2532 (1-1995) X A57126 (1-314)

Db 25 G1uargThrlLeuLeuAsnCysAArgProphofant1-nt-116 129

150 GAGGCCATTAATATCAACTGCTCCAGCTTATGAGCGAAG----- 17

[illegible]

65 AlaValLeuAlaCysLeuLeuProAspArgAlaLeuArgThrArgLeuValArgGlyGlu 84

85 Leu-----|||::|||

2 // CAAGATGTTGccCTCTCTCTTCAGAC-----TGTTTTCTCACTGACTTCTG 324

325 GGTAACCTGGAGAGA GCTTCAACTCTGTTCCCGCCTTCATCAT

110 MISGLUINTRGALAGLYPROTHRAlaValTyrPheLeuArgGlyGlyPheAspGlyPhe 135

b

136 GlnGlyCysCysProAspLeu[CysSerC]n1abren1-
432

-----ACCTGCAATTTCTCAGCCTTGCTTAACTGTGGCACATT---GGGCCAACCCGA 483

484 ATTCTTCCCAATCTTTATCTTGAGTGCACCCCCAACCACGAACTTC

110 IIELEUProTyrLeuPheLuGlySerCysSerHisSerSerAspLeuGlnGlyLeuGln 195

196 A l a C y s G l y I l e T h r A l a V a l l e u b e n y a l G e n t i a m i n e
603

804 CCCGAGTCTCATTTC-----CTGCGTGTGCCCTGTGAATGACAGCTTT 645
:::|||||

646 TGTGAGAAATTTTGGCCGTGGTTGACAAATGCACTGAAATTTT

225 metvalgluileseralatrpheglnglualaileglypheirleasptripvallylsasn 248

249 SerGlyGlyArgValIleuValHisCysGlnAlaSerTyrLeu
65

/86 ATGGCTACATCATGAAGGATGCATGTCTTAGATGAAGCTTACAGATTGTGAA 825
:::|||||:::|||||
:::|||||:::|||||

Polynemus pennellii 288

OY 826 GAAAAAGACCTACTATATCTCCAAACTCAATTTCTGGGCAACTCCTGGACTATAG 885
 Db 289 GlnArgArgGlyValIleSerProAsnPheserPheMetCylMetLeuLeuGlnPheGln 308
 OY 886 AAGACAGATT 894

Db ::::: 309 ThrCGIvrs1 311

RESULT 7
B57126

N;Alternate names: mitogen-induced nuclear protein-tyrosine-phosphatase PAC-1
C;Species: *Mus musculus* /house mouse

C:Accession: B57126
R:Roohan, P.T.; Davis, P.; Meek, J. J. *Journal of Molecular Evolution* 33: 115-125 (1991)
#text_change 24-Apr-1998

A/Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
A/Reference number: A57106

A:Status: preliminary
A:Molecule type: mpnn

A;Cross-references: GB:L11330
C:Superfamily: dual associated

F;180-311/Domain: WH1-type dual specificity phosphoprotein phosphatase

Alignment Scores:

Score:	409.00	Dengru:	314
Percent Similarity:	E1 10%	Matches:	107

Query Match:	113
B:	11.31%
Indels:	32

S-10-029-345A-108_COPY_538_2532 (1-1995) X B57126 (1-314)

25 G11A rothbar out out ... 129

130 GAAGCCATTATATCAACTGCTCCAAGCTTATGAAGCGAAG----- 171

172 -----
 2-----ProlineLeuLeuValGArgAlaProGlyThrProAla 64

65 AlaAlaLeuAlaCysLeuLeuProAspArgAlaLeuArgAlaArgLeuGlyArgGly" 84

85 T...:::|||:::|||||

277 CAAGATGTTGCCCTCTCTCTCTCAGACTGTTTTCTCACTGACTTCTGGGTAACCTTGAG 335

337 **ADDITIONAL**

116 HisGluMetArgGlyProThrThrValCysPheLeuArgGlyPheIleSerPhe 135

[illegible]

433 -----ACCTGCATTTCTCAGCCTTGCTTACCTGTGCCAACATT---GGGCCAACCTCG 483

[illegible]

Db	176	IleuLeuProTyrLeuTyrLeuGlySerCysAsnHisSerSerAspLeuGlnGlyLeuGln	195
Qy	544	CAGATGGGATTGGTTATGTCTAAATGCACGATACCTCTCCAAAGCCTGACTTATTC	603
Db	196	AlAcysgLyIlethrAlaValLeuAsnValSerAlaSerCysProAsn--HisPheGlu	214
Qy	604	CCCGAGTCTCATTTCCNCGCTGCCTGTGAATGACAGCTTTTGTGAGAAATTTTCCG	663
Db	215	GlyLeuPheHsrYrYrYrSerIleProValGluAspAsngInMetValGluIleSerAla	234
Qy	664	TGCTTGGACAATCAGTAGATTTCATGAGAAGCAAAAGCCTCCAAATGGATGTGTTTA	723
Db	235	TrpHeGlnGlnAlaIleSerPheIleAspSerValIysAsnSerGlyValArgValLeu	254
Qy	724	GTGCACCTGTTAGCTGGGATCTCCCGCTCCGCCACCACATCGCTATCGCTACATCATGAAG	783
Db	255	ValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCysLeuAlaTyrLeuIleGln	274
Qy	784	AGAGTGCACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAAAGAAAAAGCCTACATA	843
Db	275	SerHisArgValArgLeuAspGluAlaPheAspPheValIlysGlnArgArgGlyValIle	294
Qy	844	TCTCCAACTTCATATTTCTGGGCCAATCCTCGACATATGAGAAAGATT	894
Db	295	SerTroAsnPheSerPheMetGlyGlnLeuLeuGlnLeuGlnIleGlnThrGlnVal	311

RESULT 8
A:56947
dual specificity phosphatase (EC 3.1.3.-) HNH2 - rat
N:Alternate names: mitogen-activated protein kinase phosphatase 2
C:Species: Rattus norvegicus (Norway rat)
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-1999
C:Accession: A56947
J:Misra-Presz, A.; Rlm, C.S.; Yao, H.; Roberson, M.S.; Stork, P.J.S.
J. Biol. Chem. 270, 14587-14596, 1995
A:Title: A novel mitogen-activated protein kinase phosphatase. Structure, expression, and
A:Reference number: A56947; MUID:95301550; PMID:1782322
A:Accession: A56947
A:Status: preliminary
A:Function:
A:Molecule type: mRNA
A:Residues: 1-393 <MIS>
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VNI-type dual specificity
C:Keywords: phosphoprotein; phosphoric monoester hydrolase
E:202-333/Domain: VNI-type dual specificity phosphoprotein phosphatase homology <VNI>
E:279/Active site: Cys (phosphoserine intermediate) #status predicted
E:285/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:			
Pred. No.:	1,81e-16	Length:	393
Score:	393.50	Matches:	115
Percent Similarity:	46.02%	Conservative:	64
Best Local Similarity:	29.56%	Mismatch:	153
Query Match:	10.69%	Indels:	57
DB:	2	Gaps:	13

[illegible][illegible]

RESULT 9
138890
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 5 - human
N:Alternate names: dual specificity phosphatase B23; dual-specificity phosphatase hvh-3;
C:Species: Homo sapiens (man)
C>Date: 16-Feb-1996 #sequence__revision 01-Mar-1996 #text_change 11-Jun-1999
C:Accession: I38890; A55313
R:Kwak, S.P.; Dixon, J.E.

J Biol. Chem. 270, 1156-1160, 1995

A>Title: Multiple dual specificity protein tyrosine phosphatases are expressed and regulated by growth factors

A'Reference number: A55332; MUID:95138103; PMID:7836374

A'Accession: J38890

A'Molecule type: mRNA

A'Residues: 1-384 <RES>

A'CROSS-references: EMBL:U01696; NID:G642012; PID:NAB06261.1; PID:G642013

A'Experimental source: Placenta

Rishibashi, T.; Bottaro, D.P.; Michieli, P.; Kelley, C.A.; Aaronson, S.A.

J. Biol. Chem. 269, 29897-29902, 1994

A'Title: A novel dual specificity phosphatase induced by serum stimulation and heat shock

A'Reference number: A55313; MUID:95050849; PMID:7961985

A'Accession: A55313

A'Molecule type: mRNA

A'Residues: 1-8 / GHV, 12-70, 'R', 72-104, 'F', 107-362, 'RCIPIQOQSSEALMORPNPAKTMESAPQDQI

A'CROSS-references: CB:I15932; NID:9606971; PIDN:AAA64693.1; PID:9606972

A'Experimental source: mammary epithelial cells

C'Genetics:

A'Gene: GDB:DUSP5

A'CROSS-references: GDB:385447

A'Map position: 10q25-10q25

C'Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity

C'Keywords: phosphoprotein; phosphoric monoester hydrolase

F'186-317/DNA: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>

F'269/Binding site: Cys (phosphocysteine intermediate) #status predicted

Alignment Scores:

Pred. No.: 1

Score: 1.85e-14 Length: 384

Percent Similarity: 360.50 Matches: 107

Best Local Similarity: 43.67% Conservative: 62

Query Match: 27.65% Mismatches: 143

Indels: 75

Gaps: 10

US-10-029-345A-108 COPY 538 2532 (1-1995) x I38890 (1-384)

DY CTGTGGAAAGTCGAACGAAAGAAGTCTATTTATGACCAGCATTTGGGAATAC 111
||| :
15 LeuArglysgValAlaAlaAlaArgCysValValLeuAspCysArgProTyrlleuAlaIaph 34

DY AATACATCCCAACATTGTTGGAAGCCATTAAATATCAACTCTCAAGCTTATGAAAGCGAAG 171
||||| :
35 AlaAlaSerIasnValArgglySerLeuAsnValAsnLeuAsnSerValValLeuArgArg 54

DY TTGCACAACGACAAAGTGTTAATTACAGAGCTATCCAGATTCCAGCAATTAAGTT 231
:::
55 AlaArgglyValAlaVal-----SerAlaArgTyrlleuValLeu 66

DY GACATTGATTGACGTACG-----AAGGTT 255
::: :
232 GACATTGATTGACGTACG-----AAGGTT 255

DY ProAspGlnAlaAlaArgAlaArgLeuLeuGlnGlnGlyGlyValAlaAlaVal 86
67
256 GTAGTTTACGATCAAGTCGCCAAGATGTTGCTCTCTCTTCAGACGTGTTTCTCACT 315
||||| :
87 ValValAlaLeuAspGlnGlySerArgHisTrpGlnIlySleuArgGlnGlnSerAlaAlaArg 106

DY GAATCTTCGCGTAAACCTGAGAGAAGCTTCAACTCT-----GTTCACCTGCTTSCA 366
||||| :
107 ValValAlaLeuThrSerLeuAlaCysLeuProAlaGlyProArgValIlyrPheLeuIlys 126

DY GGTGGGTTGCTGACGTTCTCTGTTGTTCCCTGGCCCTCTGT-----408
127 GlyIlyTyrgIumThPheTySerGlnIlyrProGlnCysCysValaIaspValIysProIle 146

DY -----GAAGGAAAAATCCACTTACTGCTTACCTTACCTTATGCTTTCAGCTTGC 453
409
147 SerGlnGlnIyrsIleGlnSerGlnArgAlaLeuIleSerGlnCysGlyIys-----153

DY TYAACCTGTGGCAACATT-----GGGCAACCCCAATT 486
454
164 ---ProValValaIasValSerTyArgProAlaTyrlasPdnIlyGlnIlyProValGluIle 182

QY	487	CTTCCCAATCTTTATCTTGCGCTGCCAGCAGATGTCTCCAAACAGACCTGATACAGACG	546
Db	183	LeuProHeuLeuYtrLeuGlySerAlaTrpHisAlaSerLysCysGluPheLeuAlaSn	202
QY	547	AATCGAGTTGTTATATGTTAAATGCCGCTTATACCTGTCCAAAGCCTGATTATCCC	606
Db	203	LeuHisLeuThrAlaLeuLeuSerValSerArgTrpThrSerGluAla---CysMetThr	221
QY	607	GACTTCATTTTCCCTGCGTGCCTGTGCATGATGACGCTTTTGAGAAATTTTGCCGTGG	666
Db	222	HisLeuHisTyrLysAlaTrpIleProValGluIubArgSerHisThrAlaAspIleSerSerHis	241
QY	667	TTTGACAAATTCAGTATGATTTTCATTTGAGAAAGCAAAAGCCTCCAAATGAGATGTCTTAGTG	726
Db	242	PheGlnGluAlaIleAspPheIleAspCysValArgGluLysGluLysValLeuVal	261
QY	727	CACGTGTTTAGCTGGGATCTCCCGCTCCGCACACATGCTATCGCTCATCATCATGAAGAGG	786
Db	262	HisCysGluAlaGlyIleSerArgSerProThrIleCysMetAlaTrpLeuMetLysThr	281
QY	787	ATGACATGCTTTTATGATGAAAGCTTACAGATTGTGAAAGAAAAAGACTTATATCT	846
Db	282	LysGlnPheArgLeuLysGluAlaPheAspTyrIleLysGlnArgTrpSerMetValSer	301
QY	847	CCAAATTCATTTTCTCGGGCCCACTCTGACATATGAGAAAGATTAAAGCACAGACT	906
Db	302	ProAsnPheGlyPheMetGlyIleLeuLeuGlnTyrGlnSerGluIleLeuProSerThr	321
QY	907	-----GGAGCATCATGGGCCCAAGAGCAAACTCAAGCTG	939
Db	322	ProAsnProGlnProProSerCysGlnGlyGluAlaIleGlySerSerLeuIleGlyHis	341
QY	940	CTGCACCTGGAGAAAGCAAAAT-----	960
Db	342	LeuGlnIleThrLeuSerProAspMetGlnGlyAlaTyrCysThrPheProAlaSerValLeu	361
QY	961	GAACCTGTCTCCCTGCTGTCTCAGAGAGGTGAGAGAAAGCGAGCCCTCAGTCCAGCC	1020
Db	362	AlaProValProThrHisSerThrValSerGluLeuSerArgSerProValAla-----	379
QY	1021	TGTCCGCACTCTGCTCACTCA	1041
Db	380	-----ThrAlaThrSer	383

RESULT 10
 T46405
 hypothetical protein DKFZp43401321.1 - human
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 21-Jun-2002
 A:Accession: T46405
 R:Blum, H.; Baurerachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 Submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23034
 A:Accession: T46405
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-303 <AAA>
 A:Cross-references: EMBL:AL137704
 A:Experimental source: adult testis; clone DKFZp43401321
 C:Genetics:
 A:Note: DKFZp43401321.1

Alignment Scores:	
Pred. No.:	1,6e-13
Score:	345.00
Percent Similarity:	48.54%
Best Local Similarity:	32.69%
Query Match:	9.54%
DB:	2
	Gaps: 11
	Length: 303
	Matches: 101
	Conservative: 49
	Mismatches: 109
	Indels: 50

US-10-029-345A-108_COPY_538_2532 (1-1995) x T46405 (1-303)

```

Db      29 ValArgLarProserMet-AlaLeuGlyValSerGlnLeu-----41
Qy      326 GTAAAGTGGAGAAAGAGCTTCACTCTGTTCACTGCGAGTGGTTGCTGAGTTCT 385
Db      42 -----AlaGlyArgSerArgCysLeuGlySerGlnSer-GlnGlyGlyTyrGlnArgPhe 60
Qy      386 CTGCTGTTCCCTGCGCTGTCGTAAGAAATCCACTCA-----G 427
Db      60 eTserGlnTyrProGlnPheCysSerLysThrLysAlaLeuAlaAlaIleProProPro 80
Qy      428 TCCTCACTGCTATTTCTCAGCT-----TGC---TTCAGTTG 463
Db      80 aIProProSerAlaThrGlnProLeuAspLeuGlyCysSerSerCysGlyThrProLeuH 100
Qy      464 CCAACATT---GGGCAACCCGGAATTTCTCCCAATCTTTATCTTGGCTGCCAGCAGATG 520
Db      100 IsArgGlnGlyGlyProValGlnIleLeuBProPheLeuTyrLeuGlySerAlaTyrHis 120
Qy      521 TCCTCAACAAAGAGCTGATACAGAGAAATGGATGGTTATGTTGTTAAATGCCAGCTATA 580
Db      120 laAlaArgArgAspMetLeuAspAlaLeuGlyIleThrAlaLeuLeuAsnValSerSer 140
Qy      581 CTGTCCAAAGCCGCTTATATCCCGAGTCTCATTTCTGCGCTGCTGCTGATGATGACA 640
Db      140 sPyCysProAsn---HisPheGlnGlyHisTyrGlnTyrLysCysIleProValGlnAsp 159
Qy      641 GCTTTGTGAGAAATTTTCCGTGGTGGACAAATGATGATGATTTGATGAGAAAGCAA 700
Db      159 snHisLysAlaAspIleSerSerTrpPheMetGlnAlaIleGlnTyrIleAspAlaVal 179
Qy      701 AAGCTCCCAATGATGATGTTCTAGTGCACGCTTTAGCTGGATGATCCCGCTCCGCCACA 760
Db      179 ysAspCysArgGlyArgValLeuValHisCysGlnAlaGlyIleSerArgSerAlaThr 199
Qy      761 TCCTATGCGCTCATCATGATGAAGAGATGATGATGATGATGATGATGATGATGATGAT 820
Db      199 leCysLeuAlaTyrLeuMetMetLysLysArgValaLysLeuGlnGlnAlaPheGlnPhe 219
Qy      821 TGAAGAAAGAAAGAAAGCTATATATTCGCAAACTTCAATTTTGGGCGCAATCTGAGCT 880
Db      219 alLysGlnArgArgSerIleIleSerProAsnPheSerPheMetGlyGlnLeuLeuGln 239
Qy      881 ATGAGAAAGATTAAGAACCAAGCTGAGCATGAGGCGCAAGAGCAAACTCAAGCTGC 940
Db      239 heGlnSerGlnValLeuAlaThrSerCysAlaAlaGlnAlaIleSer-----254
Qy      941 TGCACTGAGAGAGCCAAATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1000
Db      255 -----ProSerGlyProLeu-----ArgGlnArgGlyL 264
Qy      1001 AGAGCGCCCTCAGTCCACCTGTGCGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060
Db      264 yThrTrpProAlaThrPro-----ThrSerGlnPheValPheSerPheP 278
Qy      1061 CCGTGCATCCCGCCAGGCTGCCAGCTGCCAGCTGCCAGCTGCCAGCTGCCAGCTGCCAG 1120
Db      278 roVal---SerValGlyValHisSerAlaProSerSerLeuProTyrLeu-----His 295
Qy      1121 GCCCGCTGTACAGGCGCTCAGT 1143
Db      295 eTProIleThrThrSerProSer 302

RESULT 11
T32494
hypothetical protein C05B10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32494
R:Geisels, C.; Wamley, P.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid C05B10.
A:Reference number: Z21178

```

```

A:Accession: T32494
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-365 <BEI>
A:Cross-references: EMBL:AF036685; PIDN:AAB88308.1; GSPDB:GND00022; CESP:C05B10.1
A:Experimental source: strain Bristol N2; clone C05B10
C:Genes:
A:Gene: CESP:C05B10.1
A:Map position: 4
A:Insertions: 22/2; 46/2; 74/3; 163/2; 204/2; 248/1; 287/2

Alignment Scores:
Pred. No.: 1,93e-10 Length: 365
Score: 294.50 Matches: 83
Percent Similarity: 45.62% Conservative: 68
Best Local Similarity: 25.08% Mismatches: 137
Query Match: 8.15% Indels: 43
DB: 2 Gaps: 8

US-10-029-345A-108_COPY_538_2532 (1-1995) x T32494 (1-365)
Qy      19 GGAATCAAAATTGTATCTAGAGAGTTGCTGCTGCTGGAAGTGAACGAAAGATG 78
Db      12 GlyGlnGlnIleSerAlaGlnGlnPheAsnArgIlePheHisGlu-----ArgAsnVal 29
Qy      79 CTGCTAATTGATGATCCCGCCATTGTGGAAATACAAATACATCCACATTTTGGAAAGCATT 138
Db      30 IleValLeuAspCysArgSer-----AsnGlyAspSerValLysArgAlaAsn 45
Qy      139 AATATCAACTGCTCCCAAGCTTATGAACGAGAGTTGCCAACAGACAAAGTGTAAATTACA 198
Db      46 ArgLeuArgLeuProAlaLeuLeuGlnArgArgLeuMetGlyLysSerMetArgLeuSer 65
Qy      199 GAGCTCATCCAGGATTCACGCAAAATGAAGTTGATGATGATGATGATGATGATGATGATGAT 258
Db      66 ThrValProAspLeuLysAspLeuAsnAsnSerProAspGlnCysProGlnValLeuLeu 85
Qy      259 GTTTAGATCAAAAGCTCCCAAGT-----GTGCTCTCTCTCT 297
Db      86 IleProGlyAspSerIleGlnAspGlnGlnLeuSerAlaAlaLeuAlaArgAsnLeuLys 105
Qy      298 TCAGACTGTTTCTCAGTCTGATCTGAGTAACTGAGAAAGAGCTTCACTCTGTTTAC 357
Db      106 SerAsnHisTyrArgHisPheValIleGlnGlnProValLysThrLeuLeuSerGlnPhe 125
Qy      358 CTGCTTGCAGAGTGGGTTTCTGCTGCTCT-----CGTTGTTTC 396
Db      126 ProThrLeuArgAspAlaAlaAspGlnAsnTrpAsnThrThrPheGlnMetAsnSerMet 145
Qy      397 CTTGGCTCTGTGAAGAAATTCACCTTAGTCCCTGATCTTCTCAGCTTGTGTTA 456
Db      146 ProGlnAlaLysSerIleGlnGln-----AlaSerSerGlyProLeuLeu 160
Qy      457 CTGTGCGCAACATTTGGG-----474
Db      161 AsnLeuAsnGlnLeuArgLeuGlnGlnLysArgGlnGlyLysSerGlnArgAlaGlnPhe 180
Qy      475 CCAACCCGAATTTCTCCCAATCTTTATCTTGGCTGCGCAGAGATGCTTCAACAGAG 534
Db      181 ProValLysLeuThrAsnPheLeuTyrLeuGlyAsnAlaGlnThrAlaLysAsnArgAsp 200
Qy      535 CTGATACAGCAGAAATGGATGTTGTTATGTTTAAATCCAGCTATACCTCTCCAAAGCT 594
Db      201 ValLeuLysLysTyrSerIleSerHisValIleAsnValThrSerAsnLeuProAsnThr 220
Qy      595 ---GACTTATCCCGCAGTCTCATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
Db      221 PheGlnGlnAspProAsnMetArgTyrLeuAlaGlyIleSerAlaAspAsnAsnIleSerHis 240
Qy      652 AAAATTTGCGCTGTGAGCAAAATGATGATTTGATGATGATGATGATGATGATGATGATGAT 711
Db      241 AsnLeuThrLysPhePheProGlnAlaIleSerPheIleAspAspAlaArgAsnAsp 260

```

```

QY 712 GGATGTTCTAGTGCACCTGTTAGTGGATCTCCGCTCCGCAACCATGCTATGCCC 771
    |||
Db 261 SerAlaCysLeuValHisCysLeuAlaGlyIleSerArgSerValThrIleCysLeuAla 280
QY 772 TACATCATGAAGAAGATGAGACATGCTCTTGAATGAGACTTACAGATTGTGGAAGAAAA 831
    |||
Db 281 TyrLeuMetLysThrGluMetCysThrLeuAspSerAlaTyrGluTrpValGlnLysArg 300
QY 832 AGACACTATATCTCCAAATCTTCAATTTCTGGGCGCAATCTCGAGCTATGAGAAGAG 891
    |||
Db 301 AsnAlaSerIleAlaIleProAsnIleHisPheMetGlyGlnLeuThrAspTyrGluLysMet 320
QY 892 ATT-----AAGAACCAAGCTGAGACATCA 915
    |||
Db 321 LeuGlyLeuAsnSerAsnArgValGlyValSer 331

RESULT 12
T18915
hypothetical protein C04F12.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18915
R:LiLoyd, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19044
A:Accession: T18915
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-272 <WIL>
A:Cross-references: EMBL:Z81461; PIDN: CAB03837.1; GSPDB: GN00019; CESP: C04F12.8
A:Experimental source: clone C04F12
C:Genetics:
A:Gene: CESP: C04F12.8
A:Map position: 1
A:Introns: 53/1; 84/3; 204/2

Alignment Scores:
Pred. No.: 4,256-07 Length: 272
Score: 239.50 Matches: 77
Percent Similarity: 47.84% Conservative: 45
Best Local Similarity: 30.20% Mismatches: 110
Query Match: 6.63% Indels: 23
DB: 2 Gaps: 8

US-10-029-345A-108_COPY_538_2532 (1-1995) x T18915 (1-272)
QY 466 AACATGAGG-----CCAAACCGAATCTTCCCAATCTTATCTTGCTGCCAG 513
    |||
Db 4 AsnValGlyValLeuGlyMetSerGlnIleAsnAspHisLeuTyrLeuSer---Gly 22
QY 514 CGAGATGCTCTCAACAAGAGCTGATAGACAGAAATGGATGGATTGTTAGTTAAATGCC 573
    |||
Db 23 AlaGlyValLeuLysProAspLysIleLysGlnArgLysIleAsnMetIleValAsnAla 42
QY 574 AGCTATACCTGCTCCAAAGCTGATTTATCCCGAGTCTCATTTCTCGCTGCTGTG 633
    |||
Db 43 ThrThrGluGluProSer---ThrTyrMetGlnGlyValAspThrMetLysIleArgIle 61
QY 634 AATGACAGCTTTTGTGAGAAAATTTTGGCTGTGGCAAAATGAGATGATTTTCATTGAG 693
    |||
Db 62 GluAspHisProTyrAlaIleArgLeuAsnGlnHisPheAspValAlaIleAspLysIleArg 81
QY 694 AAAGCAAAAGCTCCAAATGATGATGTTCTTATGCTGATGTTAGCTGGATCTCCGCTCC 753
    |||
Db 82 AsnValLysGluLysGlyGlyLysThrLeuValHisCysMetAlaGlyValSerArgSer 101
QY 754 GCCACACCTGCTATGCTGCTACATCATGAGAAGATGAGATGCTGCTTAAATGAGATTGC 813
    |||
Db 102 AlaSerLeuValMetIleTyrLeuValLysHisGlnHisMetThrLeuArgGlnAlaTyr 121
QY 814 AGATTTGGAAGAAAAAGACTTACTATATCTCCAAATCTTCAATTTCTGGGCGCAATCTC 873
    |||
Db 122 HistyTyrValLysAlaIleArgProIleIleArgProAsnValGlyPheThrLysGlnMet 141

```

```

QY 874 CTGACCTATGAGAAAGAAATTAAACAACAGATGAGATCAGGCGCAAGAGCAATCTC 933
    |||
Db 142 ValAspTyrGluLysArgLeuArgGly-----ThrAlaSerVal 154
QY 934 AAGCTGTGCACTTGAGAACCCAAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 990
    |||
Db 155 LysMetValGlnThrProGluCysAspMetProIleProAspValTyrAlaAspArgIle 174
QY 991 -----CAGAAAAGCAGACCCCTTCACTGACCTGCTGCTGCTGCTGCTGCTGCTGAG 1044
    |||
Db 175 ArgArgMetGlnIleAsnArgGluIleSerArgHisSerLeuProSerAlaGlnSerSer 194
QY 1045 GCAGCAGGACCAAGAGCCCGTGCATCCCGCAGAGGTG---CCGACCTGCTGCTGCTGCTG 1101
    |||
Db 195 AlaSerLysGlnAlaArgAlaPheSerAlaSerThrTyrArgProSerValSerAla--- 213
QY 1102 CCGTGCCTTTAAGAGACAGCCCGCTGTGATGAGCGCTCAGTGGCTGCACTGTGCGCA 1161
    |||
Db 214 ---SerSerLeuAlaThrSerSerLeuArgArgAlaTyrSerProSerSerLeuProAla 232
QY 1162 GACAGCTGGAAGAAGACAGCAATTAAGCTCAAGCCTTCTTCTCTG 1206
    |||
Db 233 SerSerLeu-----AlaLeuThrThrSerTyrSerLeu 243

RESULT 13
T21380
hypothetical protein F26A3.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21380
R:McMurray, A.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19415
A:Accession: T21380
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-226 <WIL>
A:Cross-references: EMBL:Z78419; PIDN: CAB01700.1; GSPDB: GN00019; CESP: F26A3.4
A:Experimental source: clone F26A3
C:Genetics:
A:Gene: CESP: F26A3.4
A:Map position: 1
A:Introns: 117/2; 150/3; 186/3

Alignment Scores:
Pred. No.: 9,776-07 Length: 226
Score: 233.50 Matches: 64
Percent Similarity: 49.55% Conservative: 47
Best Local Similarity: 28.57% Mismatches: 80
Query Match: 6.46% Indels: 33
DB: 2 Gaps: 6

US-10-029-345A-108_COPY_538_2532 (1-1995) x T21380 (1-226)
QY 478 ACCCGAATCTTCCCAATCTTATCTTGCTGCCAGAGATGCTCTCAACAAGAGCTG 537
    |||
Db 15 SerGlnIleValProGlyLeuPheIle---CysGlyValSerAlaLeuSerLysAspGlu 33
QY 538 ATACAGCAGAAATGAGATGTTATGTTATGTTAAATGCCACTATACCTGTCGAAAGCTGAC 597
    |||
Db 34 MetLysLysHisLysIleThrHisIleAsnAlaThrThrGluValProAsnLeuArg 53
QY 598 TTTATCCCGAGCTCATTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657
    |||
Db 54 SerLeuLysArgLysIleGlnArgThrLysLeuTyrPleuGlnAspThrProGlnThrTyrIle 73
QY 658 TTGGCGGTTGGAGCAAAATCAGATGATTTCAATGAGAAAGCAAAAGCTTCAATGAGATGT 717
    |||
Db 74 TyrProHisIleGluGlnLeuGlnSerAspGlnIleGlnAlaLeuIleAlaAspGlyGlyLys 93
QY 718 GTTCTAGTGCACCTTTAGCTGAGATCTCCGCTCCGCAACATGCTATAGGCTCAATC 777
    |||

```


Db 94 ValIeuValHicCyValaIaglyValIserArgseralIaserIleCyLeuAlaIpheLeu 113
Qy 778 ATGAAGAGATGACATGCTTTAGATGAAGCTTACGATTTTGAAAGAAAAAGCCT 837
Db 114 LeuIyterYrRgCyArGaenLeuArgIuAlaIyRhiIseuMetIySerIyArGser 133
Qy 838 ACATATCTCCAACTTCATTTCTGGCCACTCTGTGAGCTTGAAGAAGATTAAG 897
Db 134 MecValArgProAnLeuGlyPheTrpArgIleuIleAlaIyRgiIuGlnAenValIyR 153
Qy 898 AACGAGCTGAGCATGACGAGGCCAAGACAACTCAAGCTGTGACCTGAG--AG 954
Db 154 GluAsnAlaIySer-----ValArgIeuValArgAspGluAlaIagIn 167
Qy 955 CCAATGACCTGCTCCTGCTGTCTCAGAGGCTGACAGAAAGCCAGACCCCTCAGT 1014
Db 168 ProGluIleuLeuPro-----173
Qy 1015 CCAACCTGTGCGCTGTGCTACCTCAGAGCAGACAGAAAGCCCGTGCATCCGCC 1074
Db 174 -----AspValIyRLeuAsnIleAlaIle-----ProAlaArgProAla 186
Qy 1075 AGCGTGCACGCGTGCACGCGTGCAGCGTGCCTGTTAGAGAGACAGCCGCTGTACAG 1134
Db 187 Ser-----ProGluIleuAspProAsnMetIleProAspGluProArgGluArg 202
Qy 1135 GCGCTACGTGGG 1146
Db 203 ArgAsnSerGly 206
RESULT 14
T16056
hypothetical protein F13D11.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001
C/Accession: T16056
R/Fulton, L.
submitted to the EMBL Data Library, November 1995
A/Description: The sequence of C. elegans cosmid F13D11.
A/Reference number: S69020
A/Accession: T16056
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-186 <PUL>
A/Cross-references: EMBL:U40939; NID:G1072175; PID:G1072176; PIDN:AAA81700.1; CESP:F13D11.3
A/Gene: CESP:F13D11.3
A/Introns: 30/3; 57/3; 85/3; 125/3; 172/3
Alignment Scores:
Pred. No.: 2,97e-06 Length: 186
Score: 225.50 Matches: 57
Percent Similarity: 52.60% Conservative: 34
Best Local Similarity: 32.95% Mismatches: 59
Query Match: 6.24% Indels: 23
DB: 2 Gaps: 5
US-10-029-345A-108_COPY_538_2532 (1-1995) x T16056 (1-186)
Qy 478 ACCGGAATCTTCGCAATCTTTATCTT-----GGTCCAGCAGAGATGTCCTCAAC 528
Db 12 ThGlnValArgProHIsleuPheLeuAlaIglYrGlyCyS-----IleThr 27
Qy 529 AAGAGCTGATACAGCAAGATGGATTGTTATGTGTTAAATGCCACTTATACCTGTCCA 588
Db 28 ProSerIleuLeuYgInIyRAsnIleThrHIsGlyValArgCyThrIleuLeuYThr 47
Qy 589 AAGCGTACTTTATCCCGAGTCTCATTTCTGCGTGTGCTGTGGAATGACGTTTGT 648
Db 48 LysPro-----IleIySgIyLeuAspArgIleIleIuValProValaIaAspAsnThrLeu 65
Qy 649 GAGAAATTTTGGCGTGTGAGCAAAATCATGATGATTTCAATTGAGAAAGCAAGCTCC 708

Db 66 AlaIyIleThrGlnIyRPheduProValIalYyTerIleGluApeAlaIyGln 85
Qy 709 AATGATGCTTTAGTGCATGCTTTAGCTGGATCTCCCGCTCCGACCATGCTATC 768
Db 86 GlyHIsAsnThrValIleYrCyValaIaglyValIserArgseralIatIleThrIle 105
Qy 769 GCGTACATGCAAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGCAAGAA 828
Db 106 ValIyRleuMetValIthrGluAenIleSerIleGluAlaIyRleuGlnValaIagIn 125
Qy 829 AAAAGCTACTATATCTCCAACTTCATTTCTGGCCACTCTGTGACTATGAGAG 888
Db 126 ValArgProIleIleSerProAnIleGlyPheTrpArgIleuIleAlaIyRPhedu 145
Qy 889 AAGATTAAGACCAAGACTGAGCA-----TCAGGCGCAAGCAAACTCAACTG 939
Db 146 GlnArgAsnGlyAsnAlaIserValGluIleIleSerIyArgMetAlaArg----- 162
Qy 940 CTGACCTGGAGAAAGCAAAATGAACCTGCTGTGCTG 978
Db 163 -----ProValProSerVal 167
RESULT 15
T39698
protein tyrosine phosphatase - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T39698
R/Wood, V.; Skelton, J.; Churcher, C.M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, July 1999
A/Reference number: Z21870
A/Accession: T39698
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-330 <MO>
A/Cross-references: EMBL:AL109652; PIDN:CA851765.1; GSPDB:GN00067
A/Experimental source: strain 972h-; cosmid c17A3
C/Genetic:
A/Gene: p1041
A/Map position: 2
Alignment Scores:
Pred. No.: 4,05e-06 Length: 330
Score: 223.50 Matches: 53
Percent Similarity: 52.59% Conservative: 18
Best Local Similarity: 39.26% Mismatches: 63
Query Match: 6.18% Indels: 1
DB: 2 Gaps: 1
US-10-029-345A-108_COPY_538_2532 (1-1995) x T39698 (1-330)
Qy 478 ACCGGAATCTTCGCAATCTTTATCTTGGTGCAGCAGAGATGCTCAACAGAGCTG 537
Db 48 SerGluIleSerIyAsnLeuYrIleSerSerTrpYrHIsArgIuLeuValSer 67
Qy 538 ATACAGCAAGATGGATTGTTATGTGTTAAATGCCAGCTATACCTGCCAAGCCCTGAC 597
Db 68 ThrSerAspIySgIyIleAspYrThrLeuSerAlaMetSerIleAsnProAnLeuSer 87
Qy 598 TTTATCCCGAGCTCATTTCTGCGTGTGCTGTGATGACAGCTTTTGAGAAATTT 657
Db 88 ---ValProGluIleGlnIleHIsleuTrpLeuGlnIleGluAspSerIserGlnAenIle 106
Qy 658 TTGCGTGTGAGCAAAATGATGATTTGTTAGAGAAAGCAAAAGCTTCATGATGT 717
Db 107 LeuGlnIyRPheduIleuYSerIleuYRheIleAlaIleuSerIyAsnAlaIyS 126
Qy 718 GTTCTAGTGCATGTTTATGCTGGATCTCCGCTCCGCCCATCGTATCGCTCATC 777
Db 127 ValIeuValHicCySgheAlaIyIleSerArgSerValIthrIeuValaIaIyRleu 146
Qy 778 ATGAAGAGATGACATGCTTTAGATGAAGCTTACAGATTTTGAAAGAAAGAAAGCCT 837

us-10-029-345a-108_copy_538_2532.rpr

Page 12

Search completed: June 21, 2004, 12:41:49
Job time : 46.7285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:21 ; Search time 20.9747 Seconds
(without alignments)
9905.280 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 3615
Sequence: 1 atggcccatgagatgatgtg.....tggaatcatgagctctcc 1995

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xljp
-Q=/cgn21/USPTO.spool.p/US10029345/runat.21062004.122815.4177/bsp.query.fasta_1.3278
-DB=SwissProt 42 -QFMT=fastan -SUFFIX=esp -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCNALGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIZ=500 -MINLEN=50 -MAXLEN=200000000
-USBR=US10029345.QCGN.1.1.33@runat.21062004.122815.4177 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt.42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3406	94.2	665	DUSG_HUMAN	Q9by84 homo sapien
2	1326	36.7	625	DUS8_HUMAN	Q13202 homo sapien
3	1302	36.0	663	DUS8_MOUSE	O09112 mus musculu
4	515	14.2	619	VHPI_CAEEL	Q10038 caenorhabdi
5	471.5	13.0	381	DUS6_HUMAN	Q16828 homo sapien
6	469	13.0	482	DUS6_HUMAN	Q95666 homo sapien
7	468	12.9	483	DUS6_MOUSE	Q95680 mus musculu
8	466.5	12.9	381	DUS6_MOUSE	Q95681 mus musculu
9	465.5	12.9	381	DUS6_RAT	Q64346 rattus norv
10	447	12.4	394	DUS4_HUMAN	Q13115 homo sapien
11	444	12.3	375	DUS4_CHICK	Q9J171 gallus gall
12	439.5	12.2	395	DUS4_RAT	Q62767 rattus norv
13	433	12.0	367	DUS1_MOUSE	P28563 mus musculu
14	423	11.7	367	DUS1_RAT	Q64623 rattus norv
15	420	11.6	367	DUS1_HUMAN	P28562 homo sapien
16	418	11.6	314	DUS2_HUMAN	O05923 homo sapien
17	417	11.5	384	DUS9_HUMAN	Q99956 homo sapien
18	409	11.3	318	DUS2_MOUSE	Q05922 mus musculu

19	390	10.8	320	DUS7_HUMAN	Q16829 homo sapien
20	386.5	10.7	320	DUS7_MOUSE	Q91436 mus musculu
21	367	10.2	384	DUS7_RAT	O54838 rattus norv
22	361.5	10.0	280	DUS7_RAT	O63340 rattus norv
23	360.5	10.0	384	DUS6_HUMAN	Q16650 homo sapien
24	269	7.4	198	DUSE_HUMAN	Q95147 mus musculu
25	268	7.4	198	DUSE_MOUSE	O95117 mus musculu
26	205.5	5.7	188	DUS1_HUMAN	O8ne30 homo sapien
27	205.5	5.7	217	DUS1_HUMAN	O8wt42 homo sapien
28	205.5	5.7	223	STYX_HUMAN	O8wt40 homo sapien
29	205	5.7	223	STYX_MOUSE	O60959 mus musculu
30	204	5.6	278	PMPI_SCHPO	O13453 schizosacch
31	202	5.6	364	PMPI_YEAST	Q02256 saccharomyc
32	201.5	5.6	313	MSTY_HUMAN	Q9Y618 homo sapien
33	195	5.4	185	DUS3_HUMAN	P51452 homo sapien
34	195.5	5.4	295	DUSE_HUMAN	Q9h12 homo sapien
35	194.5	5.4	185	DUS3_MOUSE	Q9d73 mus musculu
36	189	5.2	489	MSG5_YEAST	P38530 saccharomyc
37	185	5.1	807	YB9T_YEAST	P38148 saccharomyc
38	183.5	5.1	276	PTP3_CHLEU	Q39491 chlamydomon
39	177	4.9	125	DUSE_MOUSE	O8r42 mus musculu
40	173.5	4.8	339	DUSC_MOUSE	Q9d02 mus musculu
41	170.5	4.7	292	YO42_CAEEL	P34650 caenorhabdi
42	163.5	4.5	340	DUSC_HUMAN	Q9un16 homo sapien
43	161	4.5	198	DUSD_HUMAN	Q9ui16 homo sapien
44	158.5	4.4	1229	P121_HUMAN	Q9Y2n3 homo sapien
45	157	4.3	209	YIL3_YEAST	P40479 saccharomyc

ALIGNMENTS

RESULT 1
DUSG_HUMAN
ID DUSG_HUMAN STANDARD; PRT; 665 AA.
AC Q9by84; Q9C0G3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase
DE phosphatase 7) (MKP-7).
GN DUSP16 OR MKP7 OR KIAA1700.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21486429; Pubmed=11489891;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.,
RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,
RT functions as a shuttle protein.";
RT J. Biol. Chem. 276:39002-39011(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21082932; Pubmed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.,
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:347-355(2000).
CC - FUNCTION: Involved in the inactivation of MAP kinases.
CC - CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC - CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC - CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC - SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC - SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC - SIMILARITY: Non-receptor class dual specificity tyrosine phosphatase.
CC - SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB052156; BAB40814.1; -
 DR EMBL; AB051487; BAB21791.1; ALT_INIT.
 DR HSSP; O16828; IMKP.
 DR Genew; HGNC:17909; DUSP16.
 DR MIM; 607175; -
 DR GO; GO:0005737; Cytoplasm; TAS.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0004721; F:protein phosphatase activity; TAS.
 DR GO; GO:0016311; P:dephosphorylation; TAS.
 DR GO; GO:0000188; P:inactivation of MAPK; TAS.
 DR GO; GO:0045209; P:leptomycin B-sensitive MAPK phosphatase nuc. .; TAS.
 DR GO; GO:0045204; P:MAPK nucleus export; TAS.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; MAPK_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHNTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANES 3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR HydroLase; Nucleic protein.
 FT DOMAIN 22 137 RHODANES.
 FT ACT_SITE 227 289 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 244 244 PHOSPHOCYSTEINE INTERMEDIATE (BY
 SQ SEQUENCE 665 AA; 73101 MW; 1BD853FF08460DF CRC64;
 SIMILARITY).

Alignment Scores:
 Pred. No.: 3, 66-206
 Score: 3406.00
 Percent Similarity: 99.85%
 Best Local Similarity: 99.70%
 Query Match: 94.22%
 DB: 1
 Gaps: 0

US-10-029-345a-108_COPY_538_2532 (1-1995) x DUSG_HUMAN (1-665)

QY 1 ATGGCCCATGAGATGTTGAACTCAATTTGTTACTGAGAGGTTGGTCTGCTGGAA 60
 Db 1 MetAlahsglmeelleglYthrglnlevalThrgluarglevalAlaleuenglu 20
 QY 61 AGTGAACGGAAAAAGTGTCTAATTTAGTACCGGCAATTTGGAAATCAATACATCC 120
 Db 21 SerGlyThngluValleuileuileaspserrpPropheValGluTYAAsnThSer 40
 QY 121 CACATTTTGAAGCCATTATATCAACTCTCAAGCTTATGAAGCGAAGTTGCAACG 180
 Db 41 HstleuengluAlaleasnleasnCyserleuileuileysargaleuengln 60
 QY 181 GACAAAGTTAAATTAAGAGCTCATCCAGATTCACCGAAACATAAGTTGACATGAT 240
 Db 61 AspIysValleuilethngluileuileghlhiserrlalyshlryValAspIleAsp 80
 QY 241 TGCACTGCAAGAGTTGATGTTAGATCAAGATCCCAAGATTTGAGCTCTCTCTTCA 300
 Db 81 CysSerGlnIysValValIyAspGlnSerSerGlnAspValAlaserleuSer 100
 QY 301 GACTGTTTCTCACTGACTTCTGGGTTAACTGGAAGAAGCTTCACTCTGTCACCTG 360
 Db 101 AspCysheleuThrValleuengluIyIysleuGlnIySerPheksnserValHstleu 120

QY 361 CTTCGACGTTGGTTTCTGAGTTCTCTGTTGTTTCCCTGGGCTCTGGAAGAAATCC 420
 Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlnIySer 140
 QY 421 ACTCTAGTCCCTACCGCAATTTCTAGGCTTGTACTCTGTTGCCAATTTGGGCAACC 480
 Db 141 ThrleuValProThrCysIleSerGlnProCysleuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATTTTCCCAATCTTTATCTTGCTGCGCCAGCAGATGCTCCCAAGAGGCTATA 540
 Db 161 ArgIleleuProAsnleuTyIleuGlyCysGlnArgAspValleuAsnIyGlnleuMet 180
 QY 541 CAGCAGAAATGGAGTTGGTTATGTTAAATCCAGCTATACCTGTCCAAAGCTGACTTT 600
 Db 181 GlnGlnAsnGlyIleGlyTyValleuAsnAlaserAsnThrCysProIyProAspPhe 200
 QY 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACGTTTGTGAGAAATTTTG 660
 Db 201 IleProGlnSerHisPheleuArgValProValAsnAspSerPheCysGlnIySileu 220
 QY 661 CCGTGTGTGACAAATGATGATTTTCATTGAGAAAGCAAAAGCTTCAATGATGTTT 720
 Db 221 ProThrleuAspIySerValAspPheIleGlnIyValIyAlalyAlaserAsnIyCysVal 240
 QY 721 CTAGTCACTGTTAGCTGGGATCTCCGCTCCGACCATGCTATGCGCTTACATCATG 780
 Db 241 LeuValHisCysleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyIleMet 260
 QY 781 AAGAGATGACATGCTTAAATGAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
 Db 261 LysArgMetCaspMetSerleuAspGlnAlaTyArgPheValIyGlnIyArgProThr 280
 QY 841 ATATCTCAACTTCAATTTCTGGGCAACTCTGACTGATGAGAAAGAAATTAAGAC 900
 Db 281 IleSerProAsnPheAsnPheleuGlyGlnleuAspIyTyGlnIySylIyAsn 300
 QY 901 CAGACTGAGACATCGAGGCGCAAGCAAAAGCTCAAGCTGTGCACTTGAGAACCAAT 960
 Db 301 GlnThrGlyAlaSerGlyProIySerIySleuIySleuHstleuGlnIyProAsn 320
 QY 961 GAACTGTGCTGCTGTCTCAAGAGGTTGACAGAAAGCGAGCGCCATGCTCACCC 1020
 Db 321 GluProValProAlaValSerGlnIyGlnIySerGlnIyProIySerProPro 340
 QY 1021 TGTGCGAAGCTGTGCTCACTCAAGAGCAGACAGCAAAAGCCCTGCAATCCCGCAGG 1080
 Db 341 CysAlaAspSerAlaThrSerGlnAlaGlyGlnArgProValHisProIaSerVal 360
 QY 1081 CCAGCGTGCAGCGTGCAGCGCTGCTGTAGAGAGCGCGCTGTGTAAGCGCTTC 1140
 Db 361 ProSerValProSerValGlnProSerleuengluAspSerProleuValGlnAlaIeu 380
 QY 1141 AGTGGGCTGCACTGTGCTCGCAGACAGGCTGAAAGACAGCAATTAAGCTCAAGGTTCTTC 1200
 Db 381 SerGlyleuHstleuSerAlaAspArgleuGlnAspSerAsnIySleuIySargSerPhe 400
 QY 1201 TCTGTGATATCAATCAGTTTCAATATTCAGCGACAGATGGCAGATCTTCAATGCTTC 1260
 Db 401 SerleuAspIleuSerValSerTyIySerAlaSerMetAlaIaserleuHstIyPhe 420
 QY 1261 TCCCATCAAGAAATGCTTTGGAATCTCAAACTTCCACTACTCTGATGGAGCAAC 1320
 Db 421 SerSerSerGlnAspAlaIeuGlnIyTyIyIyProSerThrThreleuAspGlyThrsn 440
 QY 1321 AAGCTATGCAAGTTTCCCTGTTCAGAGAACTATGGCAGACATCTCCGAAACAGTCTT 1380
 Db 441 LysleuCysGlnPheSerProValGlnGlnleuSerGlnGlnThrProGlnThrSerPro 460
 QY 1381 GATAAGAGAGAGCCAGCATCTCCCAAGAGCTGACAGCCGACGCTTTCAGAGAGCAG 1440
 Db 461 AspIySglIuAlaIaserIleProIySylSleuGlnThrAlaArgProSerAspSerGln 480
 QY 1441 AGCAAGCATTCATCTCGTCAAGAACAGACAGCTGACCGCCAGAGGTCCTTTTA 1500

Db	Seq	Accession	Length
Db	481	SerIysaHsArgLeuHisSerValArgThrSerSerSerGlyThrAlaArgSerLeuIleu	500
Qy	1501	TCCTCACTGCATCGAAGTGGAGCGCTGGAGACAATTACACACACAGCTTCCTTTGGC	1560
Db	501	SerProLeuHisArgSerGlySerValGluIlePantyrHisThrSerPheLeuPheGly	520
Qy	1561	CTTTCCACACAGCCAGACCACTTCACACAAAGTCGTGTGGCTGGGCTTAAGGCTGGAC	1620
Db	521	LeuSerThrSerGlnGlnHisIleuThrIysSerAlaGlyLeuGlyLeuIleGlyTyrHis	540
Qy	1621	TCGGATATCTTGGAGCCCGCCAGACCTTACCCCTTCCTGACACAGAGCTGGATTTTCC	1680
Db	541	SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla	560
Qy	1681	ACAGAGTCTTCACACTTCTACTCTGCTCCCTCAAGCATCTACGAGAGCAGTCCAGTTACTCT	1740
Db	561	ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer	580
Qy	1741	GCCTACAGCTGCAGCCAGCTGCGCCACTTGGCGAGACCAAGTCATATTCTGTGTGGACGGG	1800
Db	581	AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg	600
Qy	1801	CAGAAAGCCAAAGTCAGAGCTGACTCGGCGGCGAGCTGGCATGGAAGAGAGCCCTTTGAA	1860
Db	601	GlnIysProSerAspArgAlaAspSerArgAspSerTyrPheIleGlnIleSerProPheGlu	620
Qy	1861	AAGCAGTTTAAACGCAGAAAGCTGCCAAATTGGATTGGAGAGCATCATGTACAGAAAC	1920
Db	621	LysGlnPheIysPheArgArgSerCysGlnMetGluPheGlyGlnSerIleMetSerGluIleu	640
Qy	1921	AGGTACCGGGAAGAGCTGGGGGAAAGTGAGGACGTCTAGCTTTTCGGGCGACATGGAA	1980
Db	641	ArgSerArgGlnGlnLeuGlyIleValGlySerGlnSerSerPheSerGlySerMetGlu	660
Qy	1981	ATCATTTAGAGTCTCC 1995	
Db	661	IleIleGlnValSer 665	
RESULT 2			
DUS8_HUMAN			
ID	DUS8_HUMAN	STANDARD;	PRT; 625 AA.
AC	Q12072;		
DT	15-JUL-1999 (Rel. 38, Created)		
DT	15-JUL-1999 (Rel. 38, Last sequence update)		
DT	15-MAR-2004 (Rel. 43, Last annotation update)		
DE	Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)		
DE	(Dual specificity protein phosphatase hnh-5).		
GN	DUSP8 OR VHS.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OC	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Fetal brain;		
RX	MEDLINE=96009533; PubMed=7561881;		
RA	Martelli K.J., Seesholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.,		
RT	hnh-5: a protein tyrosine phosphatase abundant in brain that		
RT	inactivates mitogen-activated protein kinase.";		
RL	J. Neurochem. 65:1823-1833(1995).		
CC	-1- FUNCTION: This protein shows both activity toward tyrosine-protein		
CC	phosphate as well as with serine/threonine-protein phosphate (By		
CC	similarity).		
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein		
CC	tyrosine + phosphate.		
CC	-1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +		
CC	phosphate.		
CC	-1- TISSUE SPECIFICITY: Abundant in brain, heart and skeletal muscle.		
CC	-1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.		
CC	Non-receptor class dual specificity subfamily.		
CC	-1- SIMILARITY: Contains 1 rhodanase domain.		
CC	-----		

[illegible]

QY 418 ---TCCAGTCTAGTCCCTTACCTGCAATTTCTCAGGCTTGTCTTACCTGTGTGCAATTTGGG 474
 Db 141 ProAlaAlaLeuLeuProMetSerLeuSerGlnProCysLeuProValProSerValGly 160
 QY 475 CCAACCCGAATTTCTTCCCAATCTTTATCTTGGCGCCAGGAGATGCTCTCAACAAGAG 534
 Db 161 LeuThrArgGileLeuProHisLeuGlyLeuGlySerGlnLysAspValLeuLysAsp 180
 QY 535 CTGATPACGAGAAATGGGATTTGGTATGTGTAAATGCGAGTATCTGTGCAAGGCT 594
 Db 181 LeuThrGlnGlnGlnGlyLeuSerValLeuAsnAlaSerAsnSerCysProLysPro 200
 QY 595 GACTTTATCCCGAGTCTCATTTCTGCGGTGCGCTGCAATGACAGCTTTTGGAGAA 654
 Db 201 AspPheIleCysGlnSerArgPheMetArgValProIleAsnAspAsnLysCysGlyLys 220
 QY 655 ATTTGGCGGTGGTGGCAAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCTTCAATGGA 714
 Db 221 LeuLeuProThrLeuAspLysSerIleGlnPheIleAspLysAlaLysLeuSerCys 240
 QY 715 TGTGTTCTAGTCACTGTTAGTGGGATCTCCGCTCCGACCATGCTATGCTTAC 774
 Db 241 GlnValIleValIleCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaLys 260
 QY 775 ATCATGGAAGAGATGAGCATGTCTTTAGATGAAGCTTACAGATTGTGAAAGAAAG 834
 Db 261 IleMetLysThrMetGlyMetSerSerAspAspAlaLysArgPheValLysAspArg 280
 QY 835 CCTATATATCTCAAACTTCAATTTCTGCGGCAACTCTGCACTATGAGAAAGATT 894
 Db 281 ProSerIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuGlnLysArgLysThrLeu 300
 QY 895 AAGAACGACATGAGGATCAAGGCCAAAGAGCAAACTCAAGCTGTGACCTGAGAG 954
 Db 301 LysLeuLeuAlaAlaLeuGlnGlyAspProGly-----Thr 312
 QY 955 CCAAT-----GAACCTGTCCCTGCTGTCTCAGAGGTGAGACAGAAAGCAAGAG 1005
 Db 313 ProSerGlyThrProGlnProProProSerProAlaAlaGlyAlaProLeuProArgLeu 332
 QY 1006 CCCCTGAGTCCACCTGTGCGCACTGTCTAC-----TCAGAGCGAGAGCAACAAG 1059
 Db 333 Pro-----ProProThrSerGlnSerAlaAlaThrGlyAlaAlaAlaArgGlnGly 350
 QY 1060 CCGGTGCATCCCGCAGCGTCCAGCGTCCAGCGTCCAGCGTCCGCTGTAGAGAGC 1119
 Db 351 GlyLeuSerAlaGlyGlyGlnProProAlaProProThrProProAla-----Thr 367
 QY 1120 AGCCCGCTGGTACAGCGCTCAGTGGGCTGACCTGTCCGAGACAGAGCTGAGAGAGC 1179
 Db 368 SerAlaLeuGlnGlnLysLeuArgGlyLeuHisLeuSerSerAspArgLeuGlnAspThr 387
 QY 1180 AATAAGCTCAAGCGTCTCTCTCGATGATCAATCAATGTTCAATTCAGCCAGCATG 1239
 Db 388 AsnArgLeuLysArgSerPheSerLeuAspIleLysSerAla----- 401
 QY 1240 GCAGATCTCTTACATGGCTTCTCTCATCAGAAAGTCTTGAATCTCAAACTTCC 1299
 Db 402 -----TyrAlaProSer 405
 QY 1300 ACTACTGTGATGGAGCCAC-----AAGCTATGCCAGTTTC 1335
 Db 406 ArgArgProAspGlyProGlyProProAspProGlyGlnAlaProLysLeuCysLysLeu 425
 QY 1336 ---TCCCGCT-----GTTCAAGAACTATCGAGAGCACTCCCAACCACTGCTGAT 1383
 Db 426 AspSerProSerGlyAlaAlaLeuGlyLeuSerSerProSerProAsp---SerProAsp 444
 QY 1384 AAGGAGAGCAAGCATCCCAAGAGCTGACAGCCGACGCTTCAGACAGCCAGAGC 1443
 Db 445 AlaAlaProGlnAlaAlaArgProArgProArgArgArgPro----- 459

QY 1444 AAGCATTTGATTCGGTCAGAAACGACAGCACTGGCACCCGCCAGAGTCTCTTATCT 1503
 Db 460 -----ProAlaGlySerProAlaArg-----Ser 467
 QY 1504 CCACTGCATCGAAGTGGAGCGGTGAGAGCAATTCACACACAGCTTCTTTGGGC--- 1560
 Db 468 ProAlaHisSerLeuGly-----LeuAsnPheGlyAsp 478
 QY 1561 -----CTTCCACGACGACGACGACCTCAAGAACTCTGTGCTGCGCTTGAAGGC 1614
 Db 479 AlaAlaArgGlnThrProArgHisGlyLeuSerAlaLeuSerAlaProGlyLeuProGly 498
 QY 1615 -----TGGCACTCGATATCTTTGGCCCGCCAGACTCT 1647
 Db 499 ProGlyGlnProAlaGlyProGlyAlaThrAlaProProLeuAsnSerPro-----Gly 516
 QY 1648 ACCCTTCCCTGACGACGAGCTGATTTTCCACAGATCTTCACTTACTCTGCC 1707
 Db 517 ThrProSerProAspGlyProThrCysPheSerProGlu----- 529
 QY 1708 TCAGCATCTACGAGGACGACGACGATTCCTGCTTACAGCTGACGACGCTGCCACT 1767
 Db 530 GlyAlaGlnGlyAlaGlyGlyValLeuPheAlaProPheGlyArgAlaGlyAlaProGly 549
 QY 1768 TGCAGACCAAGTCTATCTGTGCGGAGCGGAGACCAAGTACAGAGCTGACTCG 1827
 Db 550 ProGlyGlyLysLeuAspLeuArgArgArgGlnAlaAlaArgAlaGlnProArgAspAla 569
 QY 1828 CCGGAGAGCTGCGATGAGAGACCCCTTTGAAGAGCTTAAACGAGAGAGCTGCCA 1887
 Db 570 ArgThrGlyLysProGlnGlnProAlaProGlnThrGlnPheLysArgSerCysGln 589
 QY 1888 ATGAATTTGAGAGAGCATCATGTCAGAGAACAGTCAAGG-----GAAGACTGGGAAA 1944
 Db 590 MetGlnPheGlnGlnGly-----MetValGlnGlyArgAlaArgGlyGlnGlnLeuAlaAla 608
 QY 1945 GTGGGAGTCACTTCTTGTGCGGAGCATGAAATCATTTGAGTCTCC 1995
 Db 609 LeuGlyLysGlnAlaSerPheSerGlySerValGlnValIleGlnValSer 625

RESULT 3
 DUS8_MOUSE
 ID DUS8_MOUSE STANDARD, PRT, 663 AA.
 AC 009112;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 8 (EC 3.1.1.3.48) (EC 3.1.3.16)
 GN DUSP8 OR NTPP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96311565; PubMed=8731137;
 RA Theodosios A.M., Rodriguez N.R., Nesbitt M.A., Ambrose H.J.,
 RA Paterson H., McEllan-Arnold E., Boyd Y., Leversha M.A., Owen N.,
 RA Blake D.J., Ashworth A., Davies K.E.;
 RT "A member of the MAP kinase phosphatase gene family in mouse
 RL Hum. Mol. Genet. 5:675-684(1996).
 CC -FUNCTION: This protein shows both activity toward tyrosine-protein
 CC phosphate as well as with serine/threonine-protein phosphate (By
 CC similarity).
 CC -CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = protein +
 CC phosphate.
 CC -SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
 CC -TISSUE SPECIFICITY: Expressed predominantly in brain and lung.

```

-1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95518; CAA64772.1; -.
DR HSSP; Q16828; IMKP.
DR MCD; MGI:106626; Dusp8.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR00387; Tyr_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS0206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS0054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolyase, Nuclear protein.
FT DOMAIN 23 138 RHODANES.
FT FT 162 432 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 452 459 POLY-ARG.
FT FT 555 558 POLY-SER.
FT DOMAIN 559 576 POLY-GLY.
FT FT 577 600 POLY-SER.
FT DOMAIN 311 552 PRO-HICH.
FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT FT SIMILARITY).
SQ SEQUENCE 663 AA; 68847 MW; 416F429A12C1FA7C CRC64;

Alignment Scores:
Pred. No.: 2.89e-74 Length: 663
Score: 1302.00 Matches: 318
Percent Similarity: 57.38% Conservative: 102
Best local Similarity: 43.44% Mismatches: 136
Query Match: 36.02% Indels: 176
DB: 1 Gaps: 26

US-10-029-345A-108_COPY_538_2532 (1-1995) x DUS8_MOUSE (1-663)
QY 1 ATGGCCCATGAGATGATTGGAACCTCAATTGTT--ACTGAGAGGTGGTGCTGCTG 57
Db 1 MetAlaGlyAphArgLeuProArgIySValMetAspAlaIySValLeuAlaSerIeu 20
Y 58 GAAAGTGAACGCGAAAAAGTGTCTGCTTAATTGATGATCCGCGCCATTGTGGAATACA 117
Db 21 ArgGlyGlyProGlyIyProLeuValIleAspSerArgSerPheValGlyIyArgSer 40
Y 118 TCCCAACATTTTGAAGCCATTAAATCAACTGCTCCAGCTTATGAAGAGAGTTGCA 177
Db 41 CyehIleValIleuSerValAsnIleCySbSerIySValIySValIySValIySVal 60
Y 178 CAGCAACAAGTGTATTATACAGAGCTCATCCGACATTCAGCGAACAATTAAGTTGACAT 237
Db 61 GInglyIySValThrIleAlaGlyIleuIleGlnProAlaThrArgSerGlnValAspAla 80
Y 238 GATTSCAGTCAGAGGTTGATTACATCAAGTCCCAAGATGTTGCTCTCTCT 297
Db 81 ThrGluProGlnAspValValValIyArgPginSerThrArgAspAlaSerValIleuAla 100
Y 298 TCAGACTGTTTTCACAGTACTTCGCGGTAAACGAGAAAGAGTTCACACTCTGTAC 357
Db 101 AlaAspSerPheIleuSerIleIleuIleuSerIySValAspGlyCyAPheAspSerValAla 120

```

OY	358	CTGCTGACGAGGTGGAGTTGTCGAACTTCTCTGCTGTTTCTCTGCGACCTCTGGAAGAAA	417
Db	121	TTleuthrtrgLygIyPheIatrhPheserSerCySpherProGlyLeuCySgIucllyys	140
OY	418	TCGACTAGTCCCTACC--TGCAATTCACGCTTGCTTACTCTGTTGCCAATTGGG	474
Db	141	ProIatrhLeuProSerMetSerLeuSerGlnProCySleuProValProSerValGly	160
OY	475	CCAAACCGGAATTCCTCCCAATCTTTATCTTGGCTGCGCCAGCGAGATGTCTCAACAGAG	534
Db	161	LeuthrtrgIleuProHlaSleuTytleuGlySerGlnySbaPvalLeuSnllySbaP	180
OY	535	CTGATPACGCGAGATGGGATGTGGTTATGTGTTAATGCGACGTAATTCCTGCGCAAGGCT	594
Db	181	LeuethrtrgIlnaSnGlyIleSerTyValLeuSnllySbaPamSerCySProlyPro	200
OY	595	GACTTATATCCCGAGTCTCATTTCTCGTGCTGCGCTGATGACAGCTTTGTGAGAAA	654
Db	201	AspheIleCySgIuSerArgPhemeChargIleProIleSbaPamTyValCySgIuIys	220
OY	655	ATTTTGCCGTGGTTGACAAATCAGTAGATTTCATTGAGAAAACAAAAGCTCCAAATGGA	714
Db	221	LeuethrProIleuSbaPlySerIleGluPheIleSbaPlySbaTySleuSerSerCyS	240
OY	715	TGTGTTTCAGAGCACTGTTTAGCTGGGATCTCCGCTCCGCGACCATCGATGCGCTAC	774
Db	241	GlnValIleValHlaSlySleuValGlyIleSerArgSerIatrhIleValIleValTy	260
OY	775	ATCATGGAAGAGATGACATGTCTTTAGATGAACCTTACAGATTTGTGAAGAAAAAGA	834
Db	261	ITleuethrtrhMetGlyMetSerSerSbaPamSbaTyValArgPheValIySbaPArg	280
OY	835	CTTACTATATCTCCAAATTCGAATTTTCTGGGCAACTCTCTGACATAGACAAAGATT	894
Db	281	ProserIleSerProSbaPamPheSbaPheLeuGlyGlnLeuLeuGlnTyValArgSerIeu	300
OY	895	AAGAACCAGACTGSA-----GCATCAGGGCGCAAGAGCAAACTCAAGCTGCTGACTG	948
Db	301	LysLeuLeuValIaIaLeuGlnTrhSbaPlyPro-----HisIeu	313
OY	949	GAGAAAGCAAAATGAACCTGTCCCTGCTGTCTCAGAGGGTGACAGAAAAGCGAACGCC	1008
Db	314	GlythrPro--GluProIleuMetGlyProIaIaIaGlyIleProIeuProArgLeuPro	332
OY	1009	CTCAGTCACCCCTTGCCGACTCTGTACC-----TCAGAGGAGCA-----	1050
Db	333	-----ProserTrhSerGluSerIaIaIaTrhGlySerGlnIaIaIaTrhIaIaIaArg	350
OY	1051	-----GACAAAGCCCGTCGATCCCCGACGCTCCACGCTG	1089
Db	351	GlnGlySerProSerIaIaGlyIySbaPAlaProIle--ProserTrhIaIaProIaIaTrh	369
OY	1090	CCGACCGTGCAGCGCGTGTGTTAGAGAGACGCCGCTGTAACAGGGGCTCAGTGGCGTG	1149
Db	370	SerIaIeugIn-----GlnGlyLeuArgGlyLeu	379
OY	1150	CACCTGTCCGACAGAGGCTGGAAGACAGCAATAGTCAAGCTTCTCTCTCTGTGAT	1209
Db	380	HisIeuSerSerIaPArgLeuGlnaPbTrhSbaPArgLeuIySbaGysPheSerLeuSbaP	399
OY	1210	ATCAATATAGTTTCATATTCACGACGAGATGGACGATCTTACATGGCTTCTCTCATCA	1269
Db	400	ITleuSer---AlaTyValaProSerArgIaPProSbaPheProGlyProProSbaPPro	418
OY	1270	GAAATATGCTTTGAAATACTCAAAACCTTCCACTACTCTGATGGGACCAACAAGCTATGC	1329
Db	419	GlyGlnIaIaProIySleuCySlySleuSbaPserProSerGlyGlyTrhLeuGlyLeu--	437
OY	1330	CAGTTCTCCCTGTTCAGAACTATCGAGACAGACTCCGAAACCACTCTCGATMAAGAG	1389
Db	438	-----ProserProSerProIaPserPro	445
OY	1390	GAAACCGACATCCCAAG-----AAGCTGACAGCCGACGAGCTTACAGACAGCAG	1440


```

Db 258 ArgProSerIleSerProAsnPhenMetCylLeuLeuLeuGluTyrGluAsnVal 277
Qy 892 ATTAAGACCAAGACTGAGCATCAGGCGCAAGCAACTCAAGCTGTCACCTGAG 951
Db 278 Leu-----IleLysAspHisValLeuAspTyrAsn 287
Qy 952 AAGCAATGAACCTGTCCTGCTGCTCAGAGGTTGACAGAAAGAGAGAGAGAGAGAG 1011
Db 288 GluAlaSerArgProHisArgHisMetAspTyrTyrGlyProSerAspLeuCysePro 306
Qy 1012 AGTCACACCTGTCGCACTGCTACTCAGAGGACAGAGAGAGAGAGAGAGAGAGAGAG 1071
Db 307 ----ProLysValProLysSerHisSerHisSerHisSerHisSerHisSerHis 325
Qy 1072 GCCAGCGTCGCGCAGCGTCGCGCAGCGTCGCGTCGTTAGAGACAGCGCGCTGTA 1131
Db 326 GluSerSerProSerSerProSerValSerLysLysSerLysLysSerLysPro 343
Qy 1132 CAGCGCTCAGTGGGCTGACCTGTCGCGCAGACAGGCTGAGAGAGAGAGAGAGAG 1191
Db 343 ----- 343
Qy 1192 CGTTCTCTCTCTGATATCAATCAGTTTCATATTCAGCGCAGATGCGATCCTTA 1251
Db 344 -----GluThrSerSerHisLysLysSerHisLysLysSerHisLysLys 351
Qy 1252 CATGCTCTCTCTGATATCAATCAGTTTCATATTCAGCGCAGATGCGATCCTGAT 1311
Db 352 ----SerSerSerHisLysLysSerHisLysLysSerHisLysLysSerHisLys 368
Qy 1312 GGGACCAACAGAGTATGCGAGTCTCCCTGTTGAGAGAGATGAGAGAGAGAGAGAG 1371
Db 369 GlyThrSer-----SerGlyThrValAsnValAsnGlyLysArgAsnMet 383
Qy 1372 ACAGAGT-----CCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1416
Db 384 ThrMetArgLeuGlyLeuProHisArgProLysValAlaLeuGlyLeuProSerHisGly 403
Qy 1417 ACCGCC-----AGGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1464
Db 404 ThrSerValAlaGluLeuProSerProSerThrGluLeuSerHisArgLysSerPhe 423
Qy 1465 ACCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1524
Db 423 YProGluAlaIleAlaProSerThr-----ProIleLeu----- 434
Qy 1525 GTGAGAGACAAATTCACACAGAGCTTCTTTCGAGCTTTCACAGCAGCAGAGAGAG 1584
Db 435 -----AsnPheThrAsnProCyseHisAsnSerProIleIleProValAlaSerSer 452
Qy 1585 ACAGAGTCTGCTGCGCTGAGCTTAAAGGCTGAGCATCGATATCTTGCGCCCGCAG 1644
Db 452 ValGluValIle----- 456
Qy 1645 TCATACCCCTTCCCTGACAGCAGAGCTGATATTTGCCACAGAGTCTTCACTTACTCT 1704
Db 457 LeuThrLeuProThrProAlaIle-----SerSerSerSerSerThr 471
Qy 1705 GCCTCAGAGCA-----TCATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1734
Db 471 rSerGluProSerPheAsnPheserSerPheGluSerSerSerSerSerSerIleVal 491
Qy 1735 TACTTCGCTTACAGCTGACAGCAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAG 1794
Db 491 IgluAsnProPhePheAlaSerThrGluValProAlaGlySerSerSerIleSerThr 511
Qy 1795 AGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1854
Db 511 oSerGly----SerGlnSerThrProAlaSerAlaSerSerSerAlaIleSerArg 530
Qy 1855 TTGAAA 1861

```

```

Db 530 GmetLys 532
RESULT 5
DUSE_HUMAN STANDARD; PRT; 361 AA.
AC Q16828; 075109; Q9BSH6;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
DE phosphatase 3) (MKP-3) (Dual specificity protein phosphatase PYST1).
DN DUSP6 OR MKP3 OR PYST1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Forebrain;
RX MEDLINE=96312959; PubMed=8670865;
RA Groom L.A., Sneddon A.A., Aleast D.R., Dowd S., Keyse S.M.;
RT "Differential regulation of the MAP, SAP and RK/p38 kinases by Pyel1,
RT a novel cytosolic dual-specificity phosphatase."
RL EMBO J. 15:3621-3632(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Liver;
RX MEDLINE=99077745; PubMed=9858808;
RA Furukawa T., Yatsuno T., Yousef E.M., Abe T., Yokoyama T.,
RA Fukuhige S., Soeda E., Hoshi M., Hayashi Y., Sunamura M., Kobari M.,
RA Horii A.;
RT "Genomic analysis of DUSP6, a dual specificity MAP kinase phosphatase,
RT in pancreatic cancer."
RL Cytogenet. Cell Genet. 82:156-159(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT VAL-114.
RC TISSUE=Colon, Kidney, Skin, and Stomach;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grove L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diatchenko L., Marziani K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Groom J.W., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywicki M.I., Skalska U., Smilun D.E.,
RA Schmech A., Schein I.E., Jones S.O.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 204-347.
RX MEDLINE=99140299; PubMed=10048930;
RA Stewart A.E., Dowd S., Keyse S.M., McDonald N.O.;
RT "Crystal structure of the MAPK phosphatase Pyst1 catalytic domain and
RT implications for regulated activation."
RL Nat. Struct. Biol. 6:174-181(1999).
CC -!- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
CC family.
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

```



```

Db 285 LysAsnGlyValLeuValHisCysLeuAlaGlyLeuSerArgSerValThrValThr 304
Qy 766 ATGCGCTACATCATGAAGAGATGACATCTTTAGATGACAGCTTACAGATTGTGAAA 825
Db 305 ValAlaTyrLeuMetGlnTyrSLeuAsnLeuSerMetCysAsnApIaTyrApIleValTyr 324
Qy 826 GAAAAAGACCTCTATATCTCCAAACTGCAATTTCTGGCCCAACTCCCTGACCTATGAG 885
Db 325 MetLysLysSerAnIleSerProAsnPhaAsnMetGlyGlnLeuLeuAspPheGln 344
Qy 886 AAGAAGATTAAGAACACGACTGAGCATGAGGCCA 921
Db 345 ArgThrLeu-----GlyLeuSerSerPro 352

RESULT 6
DUSA_HUMAN STANDARD; PRT; 482 AA.
AC G9Y6M6;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase
DE phosphatase 5) (MKP-5).
GN DUSP10 OR MKP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99321929; PubMed=10391943;
RA Tanoue T., Moriguchi T., Nishida E.;
RT "Molecular cloning and characterization of a novel dual specificity
RT phosphatase, MKP-5."
RL J. Biol. Chem. 274:19949-19956(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2065165; PubMed=10597297;
RA Theodosiou A., Smith A., Gillieron C., Arkinstall S., Ashworth A.;
RT "MKP5, a new member of the MAP kinase phosphatase family, which
RT selectively dephosphorylates stress-activated kinases."
RL Oncogene 18:6981-6988(1999).
RN [3]
RP SEQUENCE FROM N.A.
RT TISSUE=Brain, Lung, and Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusika K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toehlyuk S., Cantucci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.U., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywdinski M.I., Skalek U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Involved in the inactivation of MAP kinases. Has a
CC specificity for the MAPK1/MAPK2/MAPK3/MAPK4 subfamily.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +

```

```

CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL, AB026436, BA81668.1, -.
DR EMBL, AF179212, AAD51857.1, -.
DR EMBL, BC031405, AAH31405.1, -.
DR HSSP, Q16828, MKP.
DR Genew, HGNC:3065, DUSP10.
DR GO, GO:0005737, Cytoplasm, TAS.
DR GO, GO:0005634, Cytosol, TAS.
DR GO, GO:0004721, F-protein phosphatase activity, TAS.
DR GO, GO:0007254, P-JNK cascade, TAS.
DR GO, GO:0006470, P-protein amino acid dephosphorylation, TAS.
DR GO, GO:0006950, Response to stress, TAS.
DR InterPro, IPR000340, D8_phosphatase.
DR InterPro, IPR008343, MAP_kinase.
DR InterPro, IPR001763, Rhodanese-like.
DR InterPro, IPR000387, Tyr_phosphatase.
DR Pfam, PF00782, DSPC, 1.
DR Pfam, PF00581, Rhodanese, 1.
DR PRINTS, PR01764, MARKPHPTASE.
DR SMART, SM00195, DSPC, 1.
DR SMART, SM00450, RHOD, 1.
DR PROSITE, PS50206, RHODANASE_3, 1.
DR PROSITE, PS00383, TYR_PHOSPHATASE_1, 1.
DR PROSITE, PS50056, TYR_PHOSPHATASE_2, 1.
DR PROSITE, PS50054, TYR_PHOSPHATASE_DUAL, 1.
DR KMW, HydroLase; Nuclear protein.
DR FT, DOMAIN 168 285 RHODANASE.
DR FT, DOMAIN 384 453 PROTEIN-TYROSINE PHOSPHATASE.
DR FT, ACT_SITE 408 408 PHOSPHOCYSTEINE INTERMEDIATE (BY
DR FT, ACT_SITE 408 408 SIMILARITY).
SQ SEQUENCE 482 AA; 52642 MW; A8CB74ABF9498CD4 CRC64;

Alignment Scores:
Pred. No.: 4,56e-22 Length: 482
Percent: 469.00 Matches: 107
Score: 57.72% Conserved: 65
Best Local Similarity: 35.91% Mismatches: 96
Query Match: 12.97% Indels: 30
DB: 1 Gaps: 7

US-10-029-345A-108_COPY_538_2532 (1-1995) x DUSA_HUMAN (1-482)
Qy 79 CTGCTATGATGATGCGGCATTTGTGCAATACATATCCACATTTTGAAGCCATT 138
Db 173 ValIleIleAspCysArgProPheMetGlnTyrAsnLysSerHisIleGlnGlyAlaVal 192
Qy 139 AATATCACTGCGCC---AAGCTTATGAGCGAAGGTGCAACAGCAAGAGTTAAT 195
Db 193 HisIleAsnCysAlaAspLysIleSerArgArgLysGlnGlnGlnGlnIleVal 212
Qy 196 ACAGACTCATCCAG---CATTCAGCGAATCATTAAGTTGACATTTGACACTCAGAG 252
Db 213 LeuAspLeuIleSerCysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 232
Qy 253 GTTGTGTTAGATGATGAAGCTCCCAAGATGTGGCTCTCTCTTCAAGACTTTTCTC 312
Db 233 IleIleValTyrAspLysLeuAsnThrArgGlnGlnGlnGlnGlnGlnGlnGln 252
Qy 313 ACTGACTCTGGGTAACTGGAAGAGAGACTTCAACTGTCTTCACTGCTTGAGGAGG 372

```

Db	253	HsIlleValLeuGluSerLeuLysArgGluGlySerGluProLeuValIleuLysGlyGly	272
Qy	373	TTTGGTGAATCTCTCGTTGTTTCCCTGGCTCTGTGAAGAAA-----	417
Db	273	LeuSerSerPheLysGlnAsnHisGluAsnLeuLysAspAsnSerLeuGlnLeuGlnGlu	292
Qy	418	-----TTCACCTGACCTTACCTGATTCCT	444
Db	293	CysArgGluValGlyGlyAlaSerAlaAspSerLeuLeuProGlnProle---	311
Qy	445	CAGCTTGCTTACCT---GTGGCAACATTGGGCAACCCGAATTTCTCCAACTTTAT	501
Db	312	---ProThrThrProAspIleGluAsnAlaGluLeuThrProLeuProPheLeuPhe	330
Qy	502	CTTGCGTCGCAAGCAGATGTCTCCACACAGACGTGATACAGCAATGGAGATTGGTTAT	561
Db	331	LeuLysAsnGluGlnAspAlaGlnAspLeuAspThrMetGlnArgLeuAsnIleGlyLyr	350
Qy	562	GTGTAAATGCACGAC-----TAACTGTGCCAAAGCTGACTTATATCCC	606
Db	351	ValIleAsnValThrThrHisLeuProLeuLysThrIstYrGlyLysGlyLeuPhe-----	368
Qy	607	GAGTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTGTGCAAAATTTTGGCCGCG	666
Db	369	-----AsnIstYrAspArgLeuProAlaThrAspSerAsnLysGlnAsnLeuAlaGlnIstYr	386
Qy	667	TTGCAACAAATCAGTAAATTCATTGAGAAAGAAAGCCCTCAATGATGTGTCTATGT	726
Db	387	PheGlnGluAlaPheGluPheIleGluGlnAlaHisGlnCysGlyLysGlyLeuLeuIle	406
Qy	727	CACGTGTTAGCGGGATCTCCGCGTCCGCCACCATGGCTATCGCTCAATCATGAAGAG	786
Db	407	HisCysGlnAlaGlyLysSerIstYrSerIstYrIleValIleAlaIstYrLeuMetLysHis	426
Qy	787	ATGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAGAAAACCTACTATATCT	846
Db	427	ThrArgMetIstYrMetThrAspAlaIstYrIstYrPheValIstYrLysArgProIleIstYr	446
Qy	847	CCAACCTTCATTTTCTGGGCCAATCTCTGCACTATGAGAAGATTGAAGAC	900
Db	447	ProAsnLeuAsnPheMetGlyGlnLeuLeuGlnPheGlnLysPheAsn 464	
RESULT 7			
DUSA_MOUSE			
ID	DUSA_MOUSE	STANDARD;	PRT; 483 AA.
AC	Q9ESS0; Q9CZY9;		
DT	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)		
DE	(Mitogen-activated protein kinase phosphatase 5) (MAP kinase		
DE	phosphatase 5) (MKP-5).		
DN	DUSP10 OR MKP5.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxId:10090;		
PN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6;		
RC	MEDLINE=20515587; PubMed=11060451;		
RA	Masuda K., Shima H., Kikuchi K., Watanabe Y., Matsuda Y.;		
RT	"Expression and comparative chromosomal mapping of MKP-5 genes		
RT	DUSP10/Dusp10."		
RL	Cytogenet. Cell Genet. 90:71-74(2000).		
RN	[2]		

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RD MEDLINE=2108560; PubMed=11217851;
RE Kawai J., Shinagawa A., Shibata K.,
RF Aizawa T., Hara A., Fukunishi Y., Yoshino M., Itoh M., Ishii Y.,
RG Aizawa K., Tzawa M., Niichi x Komoto H., Adachi J., Fukuda S.,
RH RA

[illegible]

US-10-029-345A-108_COPY_538_2532 (1-1995) x DUSA MOUSE (1-483)

QY 79 CTGCTAATGATAGCCGCGCATTTGTGGAATACAAATATCCCAATTTTGAAGCATT 138
 Db 174 ValIleIleAspCyAspGProPheMeGluTyrHemLySerHsiIleGlnGlyAlaVal 193
 QY 139 AATATCAATGCTCC--AGCTTATGACGGAAGTTGACAGCAAAAGTTTAAAT 195
 Db 194 HsiIleAsmCyAlaAspLySileSerArgArgArgLeuGlnGlnGlySileHVal 213
 QY 196 ACAGAGCTATCCAG--CATTCAGCAAAACATTAAGTTGACATTTGACTGACAG 252
 Db 214 LeuAspLeuIleSerCyAspArgGluGlyLyAspSerPheLyAspArgIlePheSerLyGlu 233
 QY 253 GTTGATGATTACGATCAAAAGCTCCCAAGATGTTGCTCTCTTCCATGACTGTTTCTC 312
 Db 234 IleIleValTyrAspGluAsnThrAspGlnProSerArgValThrProSerGlnProLeu 253
 QY 313 ACTGATCTTGGAGTAACTGAGAAAGAGTTCAACTGTTGACCTGTTGACAGTGG 372
 Db 254 HsiIleValLeuGlnSerLeuLyAspGluGlyLySerGluProLeuValLeuLyGly 273
 QY 373 TTGTCGAGTCTCTGCTGTTTCCCTGCGCTCTGTGAGAGAAA----- 417
 Db 274 LeuSerSerPheLySerGlnAsnHsiGlyAsnLeuCyAspAspSerLeuGlnGlnGlu 293
 QY 418 -----TTCACCTGAGCCCTTACCTGACTTCT 444
 Db 294 CyAspArgIleValGlyGlyAlaSerAlaAlaSerSerMetLeuProGln-----Ser 311
 QY 445 CAGCCTTGTCTTACT--GTTGCAACATTGGGCAACCCGAATTTCTCCCAATCTTAT 501
 Db 312 ValProThrThrProAspIleGluAlaAsnAlaGluLeuThrProIleuProPheLeuPhe 331
 QY 502 CTGTCGCTCCAGGAGATGTCCTCCACAGAGACTATACAGCAAGATGGATTTGTTAT 561
 Db 332 LeuGlyAsnGlnAspIleAspIleGlnAspLeuAspThrMetGlnArgLeuAsnIleGlyTyr 351
 QY 562 GTGTTAATGCCAGC-----TATACCTGTCGAAGCTGACTTATATCCC 606
 Db 352 ValIleAsnValThrThrHsiLeuProLeuTyrHsiTyrGlyLyGlyLeuPhe----- 369
 QY 607 GACTCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
 Db 370 -----AsnTyrLyAspGluProAlaThrAspSerAsnLyGlnAsnLeuArgTyr 387
 QY 667 TTGGACAATCAGTATGATTTGATGAGAAAGCAAAAGCTCCATGAGATGTTCTATG 726
 Db 388 PheGlnGluAlaPheGlnPheIleGlnGluAlaHsiGlnCyGlyLyGlyLeuLeuIle 407
 QY 727 CACTGTTTGGTGGAGTCCGCGCTCCGCGCAACATGCTATGCTTACATCATGAGAGAG 786
 Db 408 HsiCyGlnAlaGlyAlaSerArgSerAlaThrIleValIleAlaTyrLeuMetLySHis 427
 QY 787 ATGACATGCTTTAGATGAGCTTACAGATTTGTGAAAAGAAAAGACCTACTATATCT 846
 Db 428 ThrArgMetThrMetThrAspAlaTyrLyPheValLySlyAspArgProIleIleSer 447
 QY 847 CCAACTTCAATTTTCTGGCGCAACTCTGCACTATGAGAAAGATTAAGAC 900
 Db 448 ProAsnLeuAsnPheMetGlyLeuLeuGluPheGluGluAsnLeuAsn 465

RESULT 8

DUS6_MOUSE STANDARD; PRT; 381 AA.
 AC Q9DBB1; Q9D7L4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
 DE phosphatase 3) (MKP-3).
 OS DUSP6 OR MKP3.
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10930;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver, and Tongue;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glisic C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Kono H., Balderelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=23388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.M., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maira M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
 CC family (By similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, restriction -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AK005062; BAB23786.1; -;
 CC EMBL; AK009131; BAB26093.1; -;
 CC EMBL; BC003869; AAH03869.1; -;
 CC HSSP; Q16828; IMKP.
 DR MGD; MGI:1914853; Dusp6.


```

Db 351 -----ArgGIaRgIySthPrRoAlaThrPro-----Thrse 362
QY 1041 AGAGGACGACGACAAAGCCCGTGCATCCGCCAGCGTCCAGCGTCCAGCGTGA 1100
Db 362 rGInPhaIvalPhaSerPheProVal---SerValGIyValHISerAlaProSerSerIe 381
QY 1101 GCGGTGCTGTTAGAGACAGCCCGCTGCTGATACAGCGCTCACT 1143
Db 381 uProTyLeu-----HISerProIleThrThrSerProSer 393

RESULT 11
DUS4_CHICK STANDARD; PRT; 375 AA.
AC Q9PWT1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
DE phosphatase-2) (MKP-2).
GN DUSP4 OR MKP2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=20379359; PubMed=10918612;
RA Fu S.-L., Waha A., Vogt P.K.;
RT Identification and characterization of genes upregulated in cells
transformed by v-Jun.
RL Oncogene 19:3537-3545(2000).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
CC and ERK2 (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: ARL67296; AAD4656.1; -.
DR HSSP: Q16828; IMKP.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; Tyr_phosphatase.
DR Pfam: PF00782; DSPc; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS: PR01764; MAPKPHPTASE.
DR SMART: SM00450; RHOD; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS00206; RHODANASE 3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE 1; 1.
DR PROSITE: PS00385; TYR_PHOSPHATASE 2; 1.
DR PROSITE: PS00504; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase; Nuclear protein.
KM DOMAIN 25 143 RHODANASE.
FT DOMAIN 178 375 PROTEIN-TYROSINE PHOSPHATASE.

```

```

FT ACT_SITE 261 261 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SQ SEQUENCE 375 AA; 41052 MW; 17929DDCC2BEEF1 CRC64;
SIMILARITY)
Alignment Scores:
Pred. No.: 1,61e-20 Length: 375
Score: 444.00 Matches: 123
Percent Similarity: 49.49% Conservative: 73
Beet Local Similarity: 31.06% Mismatches: 152
Query Match: 12.28% Indels: 48
DB: 1 Gaps: 13

US-10-029-345a-108_COPY_538_2532 (1-1995) x DUS4_CHICK (1-375)
QY 10 GAGATGATTGGAATCTGAATTGTTACTGAGAGTTGGCTGCTGCAAGTGAACG 69
Db 9 GluMeTGIuGIySerIaLeu-----ArgArGIuValGIyArgGIuGIaIaSerGIy 26
QY 70 GAAAAAGTGTGCTAATTGATAGCGGCGCATTTGTGGAATATACATCCACATTTTG 129
Db 27 GIyArGIySerIaLeuValSerGIyArGIySerIaLeuValHISerIaGIyHISerIa 46
QY 130 GAAGCATTAAATATCAACTGCTCCAAAGCTTATGACGAAAGTTGCAACGACAAAGTG 189
Db 47 GIyAlaIaLeuValnValArgGIySerIaThrIaValArgArGIy---AlaIySerIaVal 65
QY 190 TTAATTACAGAGCTATC-----CAGCATTCAGCAAAACATAGGTTGACATTGTC 243
Db 66 SerLeuGIuGIuIleLeuProIaGIuGIuGIuValaIaArgAlaIaArgAlaIaGIyLeu 85
QY 244 AGTCAGAAAGTTGATTTATGATCAATCAAGCTCCCAAGATGTTGCTCTCTTCAGAC 303
Db 86 TyTThAlaIaValaIaLeuTyArGIyArGIySerProArGIaGIuIaLeuArGIaPhe 105
QY 304 TGTTTCTCACTGACTTCTGAGTGAATCAAGAAAGAGCTTC-----AACTCTGTTCAC 357
Db 106 SerThValaIaLeuValaIaLeuArGIaIaLeuArGIaArGIaPheIaIaIaIaIa 125
QY 358 CTGCTTGACAGTGGGTTGGTGGTGGTCTGCTGTTGCTGCTGCTGCTGCTGCTGCTG 417
Db 126 LeuLeuAlaGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIy 145
QY 418 TCCACTGATGCTCCCTACCTCCGATTCCTCAGCCCT----- 450
Db 146 LyEThrLeu-----SerSerIleSerProProSerSerAlaGIuSerLeuPheLyGIy 163
QY 451 -----TGC-----TTACCTGTGGCCACATT---GGGCCAAACCGAAATTTCTCCCAAT 495
Db 164 PheSerSerGIyGIyThrProLeuHISaPheGIuGIyGIyGIyGIyGIyGIyGIyGIy 183
QY 496 CTTTATCTTGGCTGTCAGGAGATGTCCTCAACAGAGAGCTGATACAGCAAGATGGAGAT 555
Db 184 LeuTyLeuGIySerAlaTyThrIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 203
QY 556 GGTATGTGTAAATGCCAGCTATACCTGTCGCAAAACCTGATCTTATCCCGAGTCTCAT 615
Db 204 ThAlaIaLeuValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 222
QY 616 TTCCTCGTGTGCTGCTGATGATGACAGCTTTGTGAGAAATTTTGGCGTGTGACAA 675
Db 223 TyTyGIyGIyIleProValGIuIaPheValnHISyValaIaIaIaIaIaIaIaIaIa 242
QY 676 TCAGTGAATTCATGAGAAAGCAAGCAAGCTCCATATGATGATGTTTATGTCACCTGTTA 735
Db 243 AlaIleGIuTyIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 262
QY 736 GCTGGGATGCTCCGCGCCGCAACGCTATGCTATGCTATGCTATGATGAGATGACATG 795
Db 263 AlaGIyIleSerArGIySerAlaThrIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 282
QY 796 TCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTATATCTTCCAACTTC 855
Db 283 LyIleGIuGIyAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 302

```



```

QY 643 TTTTGTGAGAAAATTTTCCGCTGGTGGACAAATCACTAGATTTCATTGAGAAAGCAAA 702
Db 252 HtlylalaaplleSerSerPheMetGluLalIleIuVrylleapalavalLys 271
QY 703 GCCTCCAAATGATGTTCTAGTACATCTGTTAGTCTGGATCTCCGCTCCGACCATC 762
Db 272 AspCyargllylrgValleuValHtCyglLhlacllylleSeragSerlathlile 291
QY 763 GCTATGCTTACATCATGAAAGAGATGCTTTAGATGAACTTACAGATTG 822
Db 292 CyLeuLathlylryleuMetctlylYargValargleuLuglualahpGluPheVal 311
QY 823 AAAGAAAAGAACTTACTATATCTCCAAATCTTCTGAGCCCAACTCCGACTAT 882
Db 312 LysglLhlylrglSerlilleSerProaPheserhetGlyglInleuLglnPhe 331
QY 883 GAGAAGAGATTAAAGACAGATGAGCATGAGGCGCAAGAGCAAACTCAAGCTGCTG 942
Db 332 GlSerGlnValleuThrThrSerCysAlaLalagLualalaser 346
QY 943 CACCTGAGAGCCCAATGAACTGCTCTGCTTCCAGAGGCTGAGCAAGAAAGCAG 1002
Db 347 -----ProserGlyPro-----LeuArgGluArgGlySalatThrPro 359
QY 1003 AGCGCCCTCAGTCCACCTGTGCGCACTGCTTACCTCAAGAGCAGAGCAAGGCC 1062
Db 360 ThrPro-----TherSerlInPheSerPhePro 370
QY 1063 GTGCATCCCGCAGCGTCCAGCGTCCAGCGTCCGCTGCTTAGAGAGACAGC 1122
Db 371 Val-----SerValGlyValHtHlaLalPaserhAnleuProIryleu-----HisSer 387
QY 1123 CGCGTGTACAGGCGCTCACT 1143
Db 388 ProIleThrThrSerProSer 394

RESULT 13
DUST_MOUSE STANDARD; PRT; 367 AA.
ID DUST1_MOUSE
AC P28563;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE 3CH134) (Protein-tyrosine phosphatase ERP).
GN DUSP1 OR PTPN10 OR MKP1 OR 3CH134 OR PTPN16.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=92158357; Pubmed=1741163;
RA Charles C.H., Adler A.S., Lau L.F.;
RT "cDNA sequence of a growth factor-inducible immediate early gene and
RT characterization of its encoded protein.";
RL Oncogene 7:187-190(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93360956; Pubmed=8355678;
RA Noguichi T., Metz R., Chen L., Mattei M.-G., Carrasco D., Bravo R.;
RT "Structure, mapping, and expression of erp, a growth factor-inducible
RT gene encoding a nontransmembrane protein tyrosine phosphatase, and
RT effect of ERP on cell growth.";
RL Mol. Cell. Biol. 13:5195-5205(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary gland;
RX MEDLINE=22388257; Pubmed=12477932;

```

```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schlier G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
RA Raha S.S., Lonnellano N.A., Peters G.J., Abramson R.D., Millaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rodriguez Y.S.N., Krzywiński M.I., Skalska U., Smilins D.E.,
RA Schenker A., Schein J.E., Jones S.J.M., Maria M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=94037096; Pubmed=8221886;
RA Sun H., Charles C.H., Lau L.F., Tonks N.K.;
RT "MKP-1 (3CH134), an immediate early gene product, is a dual
RT specificity phosphatase that dephosphorylates MAP kinase in vivo.";
RL Cell 75:187-193(1993).
CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
CC kinase ERK2 on both Thr-183 and Tyr-185.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- INDUCTION: By growth factors.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X61940; CAA43944.1; -
CC EMBL: S64851; AAB27882.1; -
CC EMBL: BC006967; AAH06967.1; -
CC PIR: A54681; S24411.
CC HSSP: Q16828; IMKP.
CC MGD: MGI:105120; Dusp1.
CC InterPro: IPR000340; DS_phosphatase.
CC InterPro: IPR008343; MAPK_phosph.
CC InterPro: IPR001763; Rhodanese-like.
CC InterPro: IPR000387; Tyr_phosphatase.
CC Pfam: PF00782; DSPC; 1.
CC PRINTS: PR01764; MAPKPHPTASE.
CC SMART: SM00450; DSPC; 1.
CC SMART: SM00450; RHOD; 1.
CC PROSITE: PS50206; RHODANSE_3; 1.
CC PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
CC PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
CC PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
CC Hydrolase; Cell cycle-
FT DOMAIN 20 137 RHODANSE.
FT ACT_SITE 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE.
FT MUTAGEN 258 258 C->S: LOSS OF ACTIVITY.
SQ SEQUENCE 367 AA; 39369 MW; 50B5F90FEBBD19AB CRC64;

```

Alignment Scores:

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X68277; CAA8338.1; -.
DR EMBL; BC022463; AAH22463.1; -.
DR PIR; S29090; S29090.
DR HSSP; Q16828; IMKP.
DR Genew; HGNC:3064; DUSP1.
DR MIM; 600714; -.
DR GO; GO:0004726; F:non-membrane spanning protein tyrosine phos. .; TAS.
DR GO; GO:0006979; P:response to oxidative stress; TAS.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANESE_3; 1.
DR PROSITE; PS50383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; Cell cycle.
FT DOMAIN 20 137 RHODANESE.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE (BY
SIMILARITY).
SQ SEQUENCE 367 AA; 39297 MW; 11BD1D39A9FCD51F CRC64;

Alignment Scores:
Pred. No.: 5.14e-19 Length: 367
Score: 420.00 Matches: 113
Percent Similarity: 46.69% Conservative: 56
Best Local Similarity: 31.22% Mismatches: 159
Query Match: 11.62% Indels: 34
DB: 1 Gaps: 7

US-10-029-345A-108_COPY_538_2532 (1-1995) x DUS1_HUMAN (1-367)

QY 43 TTGGTGGCTGCTGTAAGTGAACGAGAAAGGCTGTAATTGATAGCGGCATT 102
DB 13 LeuArgAlaLeuLeuGlyGluArgAlaAlaGlnCysLeuLeuLeuAspCysArgSerPhe 32
QY 103 GTGGAATACATATACATCCCATTTTGGAGCCATTATATCACTGCTCCAGCTTATG 162
DB 33 PheAlaPheAsnAlaGlyHisIleAlaGlySerValaAsnValaArgPheSerThrIleVal 52
QY 163 AAGCGAAGTTGCCACAGCAGCAAAAGTTAATTACAGAGCTCATCCAGATTCCAGCGAA 222
DB 53 ArgArgArgAlaAlaGlyAlaMetCysLeuGluHisIleValProAsnAlaGluLeuArg 72
QY 223 CATAGATTGACATTGATTGACAGAGGAGGATTGATTAGATTACGATCAAGCTCCAGAGAT 282
DB 73 GlyArgLeuLeuAlaGlyAlaIleThrIleAlaValLeuLeuLeuAspGluArgSerAlaAla 92
QY 283 GTTGCTCTCTCTCTTCAGACTGTTTCTCACTGACTTCTGGGTAAATG-----GAG 336
DB 93 LeuAspGlyAlaAlaArgArgPglYThrLeuAlaLeuAlaAlaGlyAlaLeuCysArgGlu 112
QY 337 AAGACCTTCACTGCTGCTGAGAGTGGGTTTGAGTTGAGTTGCTGCTGTTTC 396
DB 113 AlaArgAlaAlaGlnValaPhePheLeuLysGlyGlyTyrgAlaAlaPheSerAlaSerCys 132
QY 397 CCTGGCTCTGTGAAGAAA-----TCACTCTTA 426
DB 133 ProGluLeuCysSerIleSerIleSerThrProMetGlyLeuSerLeuProLeuSerThrSer 152
QY 427 GTCCCTACCTGCAATTTTCAGCCTTGCTTACTGTTGCCACATT----- 471

DB 153 ValProAspSerAlaGluSerGlyCysSerSerCysSerThrProLeuTyraAspGlnGly 172
QY 472 GGGCGAACCCGAATTCCTCCCAATCTTATCTTGCTGCGCCAGAGATGTCTCAACAG 531
DB 173 GlyProValaGluIleLeuProPheLeuTyrlleuGlySerAlaIleThrIleAlaSerArgLys 192
QY 532 GAGCTGATACAGCAGAAATGGATGGTATGTGTTAAATCCAGCTTACTCTCCAAAG 591
DB 193 AspMetLeuAspAlaLeuGlyIleThrAlaLeuIleAsnValSerAlaAsnCysProAsn 212
QY 592 CTGACTTATCCCGAGCTCATTTCCGCGTGGCTGATGATGACAGCTTTGTGAG 651
DB 213 ---HisPheGluGlyHisTyrlleuTyrlleuSerIleProValGluAspAsnIleValAla 231
QY 652 AAAATTTGCGGTGGTGGCAAAATCAGTATGATTTCATTGAGAAAGCAAAAGCCTCCAAAT 711
DB 232 AsplleSerSerThrPheAsnGluAlaIleAspPheIleAspSerIleLysAsnAlaGly 251
QY 712 GGAATGTTCTGATGCACTGTTTGAAGTGGATCTCCGCTCCGACCATCGCTATGCGC 771
DB 252 GlyArgValPheValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCysLeuAla 271
QY 772 TACATCATGAAGAGATGACATGTCCTTTAGATGAGCTTACAGATTGTGAAAGAAA 831
DB 272 TyrlleuMetArgThrAsnArgValLysLeuAspGluAlaPheGluPheValLysGlnArg 291
QY 832 AGACCTTACTATCTCCAACTTCATTTTCTGGGCCCACTCTGACTATGAGAGAAG 891
DB 292 ArgSerIleIleSerProAsnPheSerPheMetCysGlnLeuLeuGlnPheGluSerGln 311
QY 892 ATTAAGAACCAAGACTGAGCATCAGGCCCAAAAGCAAACTCAAGCTGCTGCACTGGAG 951
DB 312 Val-----LeuAlaProHisCysSer-----Ala 319
QY 952 AAGCGAATGAACCTGTCCCTGCTGTCTCAGAGGGTGAACAGAAAGCGAGAGC----- 1005
DB 320 GluAlaGlySerProAlaMetAlaValLeuAspArgGlyThrSerThrThrValPhe 339
QY 1006 -----CCCTCAGTCCACCCCTGTGCGCACTTGCTACTCAGAGCAGCAGACAGAAAG 1059
DB 340 AsnPheProValSerIleProValHisSerThrAsnSerAlaLeuSerTyrlleuGlnSer 359
QY 1060 CCGCGTG 1065
DB 360 Profile 361

Search completed: June 21, 2004, 12:34:52
Job time : 42.9747 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: June 21, 2004, 12:28:25 ; Search time 103.154 Seconds
(without alignments)
12204.244 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 3615
Sequence: 1 atgagccatgagatgatgttg.....ttggaatcatgtgagctctcc 1995

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+npz.model -DEV=xlp
-Q/cgnt2 1/USPTO.spool.p/US10029345/runat 21062004.122816.4193/app.query.fasta_1.3278
-DB=SPREMBL.25 -QFMT=fasta -UUFFIX=rspc -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345.QCGN.1.1.168 @runat 21062004.122816.4193 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NRG_SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3399	94.0	665	4 Q96N49	Q96N49 homo sapien

2	3379.5	93.5	662	4 Q81VT8	Q81VT8 homo sapien
3	3163	87.5	616	4 Q8N5T1	Q8N5T1 homo sapien
4	3079.5	85.2	660	11 Q92OR2	Q92OR2 mus musculus
5	2954.5	81.7	677	11 Q99MG6	Q99MG6 mus musculus
6	2613	72.3	622	11 Q99MG5	Q99MG5 mus musculus
7	1455.5	40.3	355	11 Q9AG16	Q9AG16 mus musculus
8	1327	36.7	625	4 Q86SS8	Q86SS8 homo sapien
9	1304	36.1	665	11 Q7RSZ9	Q7RSZ9 mus musculus
10	1114	30.8	300	11 Q9AG15	Q9AG15 mus musculus
11	891	24.6	206	11 Q8B2E4	Q8B2E4 mus musculus
12	653	18.1	143	4 Q96Q52	Q96Q52 homo sapien
13	628.5	17.4	657	5 Q8ST19	Q8ST19 caenorhabdi
14	515	14.2	606	5 Q8ST18	Q8ST18 caenorhabdi
15	487.5	13.5	367	13 Q7S2F3	Q7S2F3 brachydanio
16	476	13.2	439	5 Q81G35	Q81G35 caenorhabdi
17	469	13.0	382	13 Q7T2L9	Q7T2L9 gallus gall
18	469	13.0	483	11 Q8R3L3	Q8R3L3 mus musculus
19	449	12.4	383	13 Q8UM48	Q8UM48 fugu rubrip
20	447	12.4	411	4 Q13649	Q13649 homo sapien
21	445	12.3	368	4 Q8NFX0	Q8NFX0 homo sapien
22	439.5	12.2	398	11 Q8BFX3	Q8BFX3 mus musculus
23	437	12.1	382	13 Q7T2L8	Q7T2L8 brachydanio
24	433.5	12.0	369	13 Q91790	Q91790 xenopus lae
25	431.5	11.9	359	13 Q90W58	Q90W58 xenopus lae
26	431	11.9	362	13 Q803B2	Q803B2 brachydanio
27	427.5	11.8	378	13 Q91663	Q91663 xenopus lae
28	424	11.7	367	11 Q64193	Q64193 rattus sp.
29	421	11.6	367	11 Q63683	Q63683 rattus norv
30	415.5	11.5	353	13 Q42253	Q42253 gallus gall
31	415	11.5	360	13 Q7ZVL8	Q7ZVL8 brachydanio
32	413	11.4	318	11 Q80ZM1	Q80ZM1 mus musculus
33	412	11.4	436	11 Q99KC2	Q99KC2 mus musculus
34	406	11.2	452	11 Q8KLS9	Q8KLS9 mus musculus
35	406	11.2	452	11 Q7TNL7	Q7TNL7 mus musculus
36	391.5	10.8	354	4 Q8N4A4	Q8N4A4 homo sapien
37	353	9.8	411	5 Q95SV1	Q95SV1 drosophila
38	353	9.8	411	5 Q9VW55	Q9VW55 drosophila
39	345	9.5	303	4 Q9NSM1	Q9NSM1 homo sapien
40	333	9.2	476	5 Q46122	Q46122 drosophila
41	328	9.1	476	5 Q9VW8	Q9VW8 drosophila
42	306	8.5	177	11 Q9CSL5	Q9CSL5 mus musculus
43	292.5	8.1	359	5 Q44128	Q44128 caenorhabdi
44	287	7.9	1045	5 Q9NKY1	Q9NKY1 drosophila
45	287	7.9	1193	5 Q8IMU8	Q8IMU8 drosophila

ALIGNMENTS

RESULT 1
ID Q96N49 PRELIMINARY; PRT; 665 AA.
AC Q96N49;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ131411.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Nagasawa M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nishigaki K., Masuo Y., Nagai K.,
RA Isegaki T.,
RA "MEDO human cDNA sequencing project.",
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055973; BAB71060.1; -

DR HSP: 016828; IMKP.
DR GO: GO:0016787; F:Hydrolase activity; IEA.
DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO: GO:0006470; F:protein amino acid dephosphorylation
DR InterPro: IPR000304; D: 3c phosphatase.
DR InterPro: IPR008343; MAPK phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1. TYR_phosphatase.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS: PR01764; MAPKPHTRASE.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PSS0206; RHODANESE_3; 1.
DR PROSITE: PSS0383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PSS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PSS0054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein; Hydrolase; DUAL; 1.
SEQUENCE 665 AAs. 7108 MW.

QY 1741 GCCTACAGCTGACGACGCTGCCACTTGTGGAGACCAAGCTTATTCTGNGCGCAGGCGG 1800
DB 581 AATySerTySerGlnLeuProthrCysGlyAspGlnAlaIlyrSerValAArgAArg 600
QY 1801 CAGAACCAAGTACAGAGCTGACTCGCGGCGAGCTGCGATGAGAGAGAGCCCTTTGAA 1860
DB 601 GlnLysProSerSerPheArgAlaAspSerArgSerTrpHisGlnGlnSerProPheGln 620
QY 1861 AAGCAGTTTAAACGACGAGAGCTGCCAATGAAATTGGAGAGAGCATGTCAGAGAAC 1920
DB 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGCTCAGCGGAAGAGCTGGGGAAGAGCGAGTCAGCTTCTTTCGGCAGCAGTGA 1980
DB 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATGAGGCTCTCC 1995
DB 661 IleIleGlnValSer 665

RESULT 2

Q81VT8 PRELIMINARY; PRT; 662 AA.
ID 081VT8
AC 081VT8;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Similar to dual specificity phosphatase 16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RC SEQUENCE FROM N.A.
RA Straussberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC042101; AAH42101.1; -
DR GO; GO:0017017; F:Protein kinase phosphatase activity; IEA.
DR InterPro; IPR006470; F:Protein amino acid dephosphorylation; IEA.
DR InterPro; IPR008340; DS phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR00387; Tyr_phosphatase.
DR Pfam; PF00782; DSpC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTAS.
DR SMART; SM00195; DSpC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00206; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 662 AA; 72818 MW; 350534EF0652B98F CRC64;

Alignment Scores:

Pred. No.: 7,09e-251 Length: 662
Score: 3379.50 Matches: 660
Percent Similarity: 99.40% Conservative: 1
Best local Similarity: 99.25% Mismatches: 3
Query Match: 93.49% Indels: 3
DB: 4 Gaps: 1

US-10-029-345a-108_copy_538_2532 (1-1995) x Q81VT8 (1-662)

QY 1 ATGGCCATGAGATGATTGGAACCAATTGTTACTGAGAGGTTGGTGGCTGTGGAA 60
DB 1 MetAlaHisGlnMetCileGlyThrGlnIleValAlaThrGlnLeuValAlaLeuGln 20
QY 61 AGTGAACGAGAAAAGTGTCTTAATTGATGACCGGCATTGTGAATACAATACATCC 120
DB 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnIlyrAanthSer 40

QY 121 CACATTTTGAAGCCATTATATCACTGTCCCAAGCTTATGAGAGGAGTTGCAACAG 180
DB 41 HisIleLeuGlnAlaIleMetIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
QY 181 GACAAAGTGTATTATTAACAGAGCTCATCAGACATTCAGCAAAACATAAGTTGACATTGAT 240
DB 61 AspLysValLeuIleThrGlnLeuIleGlnIleSerAlaIlyrHisLysValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATTATTCATCAAGTCCCAAGATTTGCTCTCTCTTCA 300
DB 81 CysSerGlnLysValAlaValAlaIlyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTCAGCTGACTGCTGGGTAACAGAGAGCTTCAACTGTCTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGTGGGTTTCTGAGTCTCTGTTGTTTCTTCGCTCTGTGAAGAAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
QY 421 ACTTATGTCCTTACCTGATTTCTCAGCTTGTCTTACCTGTTGCCAATGGGCCAACC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTATCTTGGCTGCGCAGCGAGATGCTCTCAACAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuLysLeuGlyCysGlnAspAspValLeuAsnLysGlnLeuMet 180
QY 541 CAGCAGAAATGGAGTTGTTATGTTAAATGTCAGATATACCTGTCCAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyLysValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
QY 661 CGGTGATGGAGCAAAACAGAGATTTCATTGAGAAAGCAAAAGCTCCAAATGATGTTG 720
DB 221 ProTrpLeuAspLysSerValAlaPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTATGCTACATGTTTAACTGGGATCTCCGCTCCGCCAATGCTATGCTGCTATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIlyrIleMet 260
QY 781 AAGAGATGAGACATGCTTCTTATGATGAAGCTTAAAGATTGTGAAGAAAAGACCTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaIlyrArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGACTATGAGAGAAAGATTAGAAC 900
DB 281 IleSerProAsnIleAsnIlePheLeuGlyGlnLeuLeuAspIlyrIlyrIlyrIlyrIlyr 300
QY 901 CAGACTGAGACATCAAGGCGCAAAAGCAAACTCAAGCTGTGCACTGAGAGCCCAAT 960
DB 301 GlnThrGlyAlaIleSerGlyProLysSerLysLeuLeuLeuHisLeuGlnLysProAsn 320
QY 961 GAACTGTCTCTCTGTCTGTCAGAGGTCAGAGAGAAAGCAAGACGCCCTCATCCACC 1020
DB 321 GlnProValProAlaValSerGlnGlyGlnLysSerIlyrIlyrIlyrIlyrIlyrIlyr 340
QY 1021 TGTGCGGAGCTGCTACCTCAGAGGAGAGCAAAAGCGCGTGATCCCGCAGCGTG 1080
DB 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnIlyrProValHisProAla----- 358
QY 1081 CCCAGCGTCCAGCGCTGACGCGCTGCTGTTAGAGAGACGCCGCTGTACAGCGCTC 1140
DB 359 ---SerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 377
QY 1141 AGTGGCTGACCTGTTCGCGACAGAGCTGGAGAGACAGCAATAAGCTCAAGCTTCTCTTC 1200
DB 378 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 397
QY 1201 TCTCTGATATCAAAACAGTTTATGATGAGCAGATGGCAGCATCTTACATGCGTTC 1260

```

Db      398 SerLeuAspIleIysSerValSerIyrSerAlaSerMetAlaIaSerLeuHISgIlyPhe 417
Qy      1261 TCCATCATGAGAAAGATGCTTTGGAATACTTACAAACCTTCACTACTCGATGGAGACAC 1320
Db      418 SerSerSerGluAspAlaLeuGluTyrTyrIlySerProSerThrThrLeuAspGlyThrAsn 437
Qy      1321 AAGCATGCGCATCTCCCTGTTGAGAACTTTCGAGACAGACTCCGAAACAGATCCT 1380
Db      438 LysLeuGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 457
Qy      1381 GATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db      458 AspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 477
Qy      1441 AGCAAGCATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1500
Db      478 SerIyrArgLeuHISerValAlaGlnHISerSerSerGlyThrAlaGlnArgSerLeuLeu 497
Qy      1501 TCTCCATCGATGAAAGTGGAGAGCTGGAGAGCAATTCACACAGAGCTTCTTTGGC 1560
Db      498 SerProLeuHISerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 517
Qy      1561 CTTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
Db      518 LeuSerThrSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 537
Qy      1621 TGGATATCTTGGCCCCCAAGCTCTACCCCTTCCAGCAGCAGCAGCAGCAGCAGCAGCAG 1680
Db      538 SerAspIleLeuAlaProGlnThrSerThrProSerLeuHISerSerIlyrPheAla 557
Qy      1681 ACAGAGCTCTCAACCTTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db      558 ThrGlnSerSerHISerHISerHISerHISerHISerHISerHISerHISerHISerHISer 577
Qy      1741 GCTTACAGCTCTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db      578 AlaIyrSerCysSerGlnLeuProHISerGlyAlaSerGlnValIyrSerValAlaArg 597
Qy      1801 CAGAGCAGAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860
Db      598 GlnIyrProSerAspArgAlaAspSerAspArgSerThrIlyrHISerGlnSerProPheGlu 617
Qy      1861 AAGCAGTTTAAAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Db      618 LysGlnPheIyrArgIyrSerCysGlnMetGlnPheGlnGlnSerIlyrMetSerGlnAsn 637
Qy      1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCAAGTCTTTCGAGCAGCAGTGA 1980
Db      638 ArgSerArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 657
Qy      1981 ATGATTGAGGTCTCC 1995
Db      658 ILeIleGlnValSer 662

```

```

DR      GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR      GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR      InterPro: IPR000340; DS phosphatase.
DR      InterPro: IPR008343; MAPK_phosph.
DR      Pfam: Pf00782; DSpC; 1.
DR      PRINTS: PR01764; MAPKPHPTASE.
DR      SMART: SM00195; DSpC; 1.
DR      PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR      PROSITE: PS00056; TYR_PHOSPHATASE_2; 1.
DR      PROSITE: PS00054; TYR_PHOSPHATASE_DUAL; 1.
DR      Hypothetical protein.
FT      NON TER
SQ      SEQUENCE 616 AA; 67636 MW; 2CBOB14482F2AD72 CRC64;

Alignment Scores:
Pred. No.: 2,96e-234 Length: 616
Score: 3163.00 Matches: 613
Percent Similarity: 99.84% Conservative: 2
Best Local Similarity: 99.51% Mismatches: 1
Query Match: 87.50% Indels: 0
DB: 4 Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x Q8N5T1 (1-616)
Qy      148 TGCCTCAAGCTTATGAAAGGAGGAGTTCGAAACGAGCAAAAGTGTATATTCAGAGCTCATC 207
Db      1 CysSerIyrLeuMetIyrAspArgLeuGlnGlnAspIyrValIleThrGlnLeuIle 20
Qy      208 CAGCATTCAGCGAAACATTAAGTTGACATTTGACATTTGACATTTGACATTTGACATTTGACAT 267
Db      21 GlnHISerAlaIyrHISerIyrValAlaAspCysSerGlnIyrValIyrAsp 40
Qy      268 CAAAGCTCCCAAGAGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 327
Db      41 GlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGly 60
Qy      328 AATCTGAGAGAGAGCTTCACTGTTCACCTGCTTCAGAGTGGAGTTCGAGTCTCT 387
Db      61 LysLeuGlnIyrSerPheAsnSerValHISerLeuMetAlaGlyIyrPheAlaGlnPheSer 80
Qy      388 CGTTGTTTCCCTGGGCTCTGTGAGAGAAATCCACTTACTTCTCTCTCTCTCTCTCTCTCTCT 447
Db      81 ArgCysPheProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 507
Qy      448 CTTGCTTACTCTGTTGCCAATTTGGGCCAAGCCGAATCTTCCCAATCTTATCTTGGC 507
Db      101 ProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuIyrLeuGly 120
Qy      508 TGCAGCAGAGATGCTCCCAAGAGAGCTGATACAGCAGATGGAGATGGATGGATGGATGGAT 567
Db      121 CysGlnArgAspValAlaLeuAsnIyrGlnLeuMetGlnGlnAsnGlyIleGlyTyrValLeu 140
Qy      568 AATCCAGCTTAACTGTCCAAAGCTGACTTATCCCGAGTTCATTTCTCTGCTG 627
Db      141 AsnAlaSerAsnThrCysProIyrProAspPheIleProGlnSerHISerPheLeuIyrVal 160
Qy      628 CTTGTAATGACAGCTTTTGTGAGAAATTTTCCGCTGTTGCAATCATGATTTTC 687
Db      161 ProValAsnAspSerPheCysGlnIyrIleLeuProIlyrLeuAspIyrSerValAspPhe 180
Qy      688 ATTGAGAAAGCAAAAGCTTCGAATGATGTTCTTCAAGGACCTGTTTTCAGGATCTCC 747
Db      181 ILeGlnIyrAlaIyrAlaSerAsnGlyCysValLeuValHISerCysLeuAlaGlyIleSer 200
Qy      748 CGCTCCGACCATGAGCTATGCTGCTTACATCATGAAAGAGATGACATGCTTTAGATGAA 807
Db      201 ArgSerAlaThrIleAlaIleAlaIyrIleIleIyrAspArgMetSerLeuAspGln 220
Qy      808 GCTTACAGATTTGTGAGAAAGAAAGCACTTACTATATCTCAATCTTCAATTTCTGGGC 867
Db      221 AlaIyrArgPheValIyrGlnIyrArgProThrIleSerProAsnPheAsnPheLeuGly 240

```

QY	868	CAACSTCCGAACTAATGAGAAAGAAATTAAAGACCAAGCTGGAGCACTCAAGGCGCAAGAC	927
Db	241	GlnLeuLeuAaPrgTgGlnUlySerIleUlyVaenGlnTtrGlyAlaSerClyProUlysSer	260
QY	928	AAACTCAAGCTGCTCACCTGGAGAAAGCCAAATTAAGACTGTCCCTGCTGCAGAGGAT	987
Db	261	LysUlyLeuUlyLeuLeuHnIleUenGlnUlyProVaenGlnUProAlaValaSerUlnGly	280
QY	988	GGACAGAAAAGCGAGACGCGCCCTCAGTCCACCCCTGTGCGCACTGTGCTAATCTCAAGGCA	1047
Db	281	GlyGlnUlySerGlnUThrProUlySerProProCysAlaAaSerAlaThrSerGlnAla	300
QY	1048	GCAGAGCAAAAGCCCGCTGATCCCGCCAGCCGTCGCCAGGTGCGCAAGCGTCG	1107
Db	301	AlaGlyGlnAaPrgProValHnIaProlaAaSerAlaProSerValGlnProSer	320
QY	1108	CTGTATAGAGACAGACGCCGCTGGTAAAGGGCGCTCACTGGGCTGCACCTGTGCCAAGACG	1167
Db	321	LeuLeuGlnAaPserProLeuValGlnAlaUlySerGlnUlyLeuHnIleUlySerAlaAaPrg	340
QY	1168	CTGGAGAAGCAAGATTAAGCTCAAGCGTTCCCTTCTGTGGATATCAAAATCAATTCATAT	1227
Db	341	LeuGlnAaPserAaHnUlyLeuUlyAaPrgSerPheSerLeuAaPrlIleUlySerValaSerUly	360
QY	1228	TCAGCGACAGTGGCGAGCATCCTTAACATGGCTTCTCTCATCAAGAAATGCTTTGGAAATAC	1287
Db	361	SerAlaAaSerMetAlaIaSerLeuHnIaGlyPheSerSerSerGlnAaPrlaLeuGlnUlyUy	380
QY	1288	TACAAACCTTCCACACTCTGGATGGAGCAACAAAGCTATGGCAGATTCCCGCTGTACAG	1347
Db	381	TyrUlyPProSerThrThrLeuAaPrlUlyHnAaUlyUlyCysGlnPheSerProValGln	400
QY	1348	GAACATATCGAGACAGACTCCCGAAACCAAGTCCCTGATTAAGAGAAAGCCAGCATCCCGAAG	1407
Db	401	GlnUlySerGlnUlnTtrProGlnUlyHnSerProAaPrgUlyGlnUlnAlaSerUlnIleProUlys	420
QY	1408	AAAGCTGCAGACCGCCAGGCTTCAAGACCCAGACCAAGCAATTCGATTCGGTCAAGAAC	1467
Db	421	LysLeuGlnTtrAlaIaPrgProSerAaPserGlnSerUlyAaGlnUlnIaSerValaIaPrgThr	440
QY	1468	AGACAGCATGGGACCGCCAGAGGTCCTTTATCTCCATGCTGACATCGAAATGGAGGCGG	1527
Db	441	SerAaenSerGlyThrAlaGlnAaPrgSerLeuUlySerProLeuHnIaAaPrgSerUlySerAla	460
QY	1528	GAGGCAAAATTAACACACAGCACTTCCCTTTTCGCGCTTCCACACAGCCAGACAGCACTCCAG	1587
Db	461	GlnAaPrgAaUlyHnIaThrSerPheUlyPrgUlyLeuUlySerThrSerGlnGlnHnIleUlyHn	480
QY	1588	AAAGTGTGCTGCGCTGAGGCTTAAGGAGCTGGACATCGAATATCTTGAGCCCCCAAGACTCT	1647
Db	481	LysSerAlaIaGlyLeuGlnUlyLeuUlySerGlyUlyTrpHnIaSerAaPrlIleUlnAlaProGlnThrSer	500
QY	1648	ACCCCTTCCCTGACACAGACGCTGGATTTTGGCAGAAAGTCTTCAACATTCATCAATCTCC	1707
Db	501	ThrProSerLeuThrSerSerUlyUlyUlyPheAlaIaGlnUlySerSerHnIaPheUlySerAla	520
QY	1708	TCAGCGCATTAAGAGAGCAAGTCCAGTTAATCTGTGCTAATCACTGACGTGACGCACTGCCACT	1767
Db	521	SerAlaIaIleUlyGlnUlySerHnIaSerUlySerAlaUlyUlySerCysSerGlnIleUlyProThr	540
QY	1768	TGCGGAGACCAAGTATTAATCTGTGCGGACGGCGGAGAAAGCCAAAGTGAACAGACTGACTCG	1827
Db	541	CysGlnAaPrgInValUlySerValaIaPrgAlaGlnUlyUlyProSerAaPrgAlaAaPser	560
QY	1828	CGGCGGCACTGGCAATGAAGAGCGCCCTTTGAATAAGCAATTAACGACAGACGTCGCCAA	1887
Db	561	AaPrgAaSerUlyTrpHnIaGlnUlySerProPheGlnUlyUlyGlnPheUlyAaPrgAaPserCysGln	580
QY	1888	ATGGAATTTTGAAGAGCACTCAATGCTCAGAGAAACAGGTCAACGGAAGAGCGTGGGAAAGTG	1947
Db	581	MetGlnUlyPheGlnUlySerUlnMetSerGlnAaAaPrgSerHnIaGlnUlnUlyUlySerAla	600
QY	1948	GGCAGCTCAAGTCTTAAGCTTTTCGGGACAGATGAAATCAATTGAAGTCTCC	1995

Db 601 GlySerClnSerSerPheSerGlySerMetGluIleIleGluValSer 616

|||||
|||||

RESULT 4

ID 0920R2 PRELIMINARY; PRT; 660 AA.

AC 0920R2;

DT 01-DEC-2001 (TREMBLE). 19, Created)

DT 01-DEC-2001 (TREMBLE). 19, Last sequence update)

DT 01-OCT-2003 (TREMBLE). 25, Last annotation update)

DE MAP kinase phosphatase-7.

GN DUSP16 OR 3830417M1RIK OR MKP-7.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shira H., Watanabe M., Kikuchi K.;
RT "MKP-7", a Novel Mitogen-activated Protein Kinase Phosphatase,
RT Functions as a Shuttle Protein.";
RL J. Biol. Chem. 276:39002-39011(2001).
DR EMBL; AB052157; BAB47240.1; -;
DR HSSP; O16829; IMKP
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005634; Cnucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro; IPRO00340; DS_phosphatase.
DR InterPro; IPRO008343; MAPK_phosph.
DR InterPro; IPRO01763; Rhodanese-like.
DR InterPro; IPRO00387; TYR_phosphatase.
DR Pfam; PF00782; DSPc; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SMO0195; DSPc; 1.
DR SMART; SMO0450; RHOD; 1.
DR PROSITE; PSS0206; RHODANSE_3; 1.
DR PROSITE; PSS00383; TYR_PHOSPATASE_1; 1.
DR PROSITE; PSS0056; TYR_PHOSPATASE_2; 1.
DR PROSITE; PSS0054; TYR_PHOSPATASE_DUAL; 1.
KW Hydrolase; kinase.
SQ SEQUENCE 660 AA; 72695 MW; DB609FCDADAA309 CRC64;

Alignment Scores:

Pred. No.: 7, 82e-228 Length: 660

Score: 3079, 50 Matches: 599

Percent Similarity: 94, 29% Conservative: 28

Best Local Similarity: 90, 08% Mismatches: 33

Query Match: 85, 19% Indels: 5

DB: 11 Gaps: 3

US-10-029-345A-108_COPY_538_2532 (1-1995) x Q920R2 (1-660)

QY 1 ATGGCCCATAGAKATGTATTGGAACTCAATTTGACTGAGAGGTGTGGTCTCTGTGGA 60

DB 1 MetAlaHisgLeuMetIleGlyThrGlnIleValIThrGlnSerLeuValAlaLeuenglu 20

QY 61 AGTGGAACCGAAAAAGTCGGCTTAATGTATGACCGGCACATTTTGGATATACATACC 120

DB 21 SerGlyThrGlnuysValLeuLeuIleAspSerArgProPheValGluTyrrhenThSer 40

QY 121 CACATTTGGAGACCTTAATAATCAACTGCTCCAAGCTTATGAAACCGAAGTTGCCAACG 180

DB 41 HisIleengLuIAleIleHisenIleAcnysserLySleuMetLyARGArgLeugIngin 60

QY 181 GACAAGAAGTGTATTTCACAGAGCTCATCCAGCATTTAGGAAAACATTAAGCTGACATTGAT 240

DB 61 AspIyValLeuIleuIleThrGluLeuIleGlnHisserAlaValVSHSLyValAspIleEsp 80

QY 241 TCGAGTCAGGAAGTTGATGTTACGATTAAGCTCCCAAGATGTTGCTCTCTTCA 300
 Db 81 CysaenglnlysvaValValValTyrAspGlnSerSerGlnAspValGlySerLeuSerSer 100
 QY 301 GACTGTTTCTGACTGTATCTTGGGTAACCTGAGAGAGCTTCAACTCTTTCACCTG 360
 Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnAspSerPheAsnSerValHisLeu 120
 QY 361 CTTCGAGGTGGGTTTGGTGGATCTCTCGTTGTTTCCCTGAGCTCTGTAAGAGAAATCC 420
 Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlySerSer 140
 QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCCCTTGCTTCTGTTGCAACATTTGGGCAAC 480
 Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATCTTCCCAATCTTATCTTGGCTGCGACCGAGATGTCCTCAACAGAGACTGATA 540
 Db 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnAspValLeuAsnLysAspLeuMet 180
 QY 541 CACAGAAATGGGATTTGTTATGTTAATGACCACTAATCTGCTCCCAAGCTGACTTT 600
 Db 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
 QY 601 ATCCCGAGTCTCATTTCTCGTGGTGGCTGTGAATGACAGCTTTTGTGAAATTTTG 660
 Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
 QY 661 CCGTGGTTGAGCAATTCAGTATTTCAATTGAGAAACAAAGGCTCCATTTGATGTTGT 720
 Db 221 ProThrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
 QY 721 CTAGTCAGCTGTTTATGCTGGATCTCCGCTCCGCCACATGCTTATCCCTTACATCATG 780
 Db 241 LeuIleHisCysIleAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGACTACT 840
 Db 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlnLysArgProThr 280
 QY 841 ATATCTCAAACTTCATTTTCTGGGCAACTCTCTGACTATGAGAAAGATTAAAGAAC 900
 Db 281 IleSerProAsnPheAsnPheMetGlyGlnLeuMetAspLysGlnLysThrIleAsnAsn 300
 QY 901 CAGACGTGACATCAGGGCCAAAGACCAACTCAAGCTGCTGACCTGAGAGAGCAAAAT 960
 Db 301 GlnThrGlyMetSerGlyProLysSerLysLeuLysIleLeuHisLeuAspLysProSer 320
 QY 961 GAACCTGTCTCTGCTCTCAGAGGGTGGACAGAAAGCGACGCTTCAGTCAACCC 1020
 Db 321 GluProValProAlaAlaSerGlnGlyIleTyrLysSerAlaLeuSerLeuSerProPro 340
 QY 1021 TGTGCGCATCTGCTTACTCAGAGGACAGACAGCAAGGCGCTGACATCCCGCAGCGTG 1080
 Db 341 CysAlaAsnSer---ThrsrGlnAlaSerGlyGlnArgLeuValHisProAla----- 357
 QY 1081 CCCAGGTGCCAGGTCGAGCGGTCGCTGTTAAGACAGCCCGCTGTGACAGGCGCTC 1140
 Db 358 ---SerValProSerLeuGlnProSerLeuGlnAspSerProLeuValGlnAlaLeu 376
 QY 1141 AGTGGGTCGACCTGTCCGACAGACAGGCTGAGACAGCAATTAAGCTCAAGGCTTCCTTC 1200
 Db 377 SerGlyLeuGlnLeuSerSerGlnLysLeuGlnAspSerThrLysLeuLysAspSerPhe 396
 QY 1201 TCTGTGATATCAATCAGTTTATATTCAGCGACAGATGGCAGACTCTTACATGGCTTC 1260
 Db 397 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 416
 QY 1261 TCCCATCAAGAAATGCTTTGGAATTAACAACCTTCACTGATCTGATGGGACCAAC 1320
 Db 417 ---SerSerGlnGlnAlaLeuAspLysTyrCysLysProSerAlaThrLeuAspIleThrAsn 435
 QY 1321 AAGCTATGCAAGTTCTCCCTGTTTCAGGAACATATGGAGACAGATCTCCGAAACAGTCT 1380

Db 436 LysLeuCysGlnPheSerProValGlnLysValSerGlnAsnProGlnThrSerPro 455
 QY 1381 GATAGAGAGAGAGAGAGAGATCCCAAGAGCTGACAGACCGGACGCTTCAGACAGCAG 1440
 Db 456 AspLysGlnGlnAlaHisIleProLysGlnProGlnProAlaProAspSerGlnSerGln 475
 QY 1441 AGCAAGCATTCATTCGATTCGATGAGAAGCAGACAGAGCTGGACCGCCAGAGTCCCTTTTA 1500
 Db 476 ValThrArgLeuHisSerValArgThrGlySerSerIleGlySerThrGlnArgProPhePhe 495
 QY 1501 TCTCAGCATTCGAAAGTGGAGAGGCTGAGAGACAAATTAACACACAGCTTCTTTTGGC 1560
 Db 496 SerProLeuHisAspSerGlySerValGlnAspAsnTyrHisThrAsnPheLeuPheGly 515
 QY 1561 CTTTCACAGCAGCAGCAGCAGCTCAAGATCTGCTGGGCTTGAAGGCTGGCAC 1620
 Db 516 LeuSerThrSerGlnGlnArgLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrHis 535
 QY 1621 TCGATATCTTGGGCCCCCGACCTTACCCCTTCCCTGACAGCAGCTGTGATTTTGGC 1680
 Db 536 SerAspIleLeuAlaProGlnSerSerAlaProSerLeuThrSerSerTyrPheAla 555
 QY 1681 ACAGATCTCTACACTTCTACTCTGCTCAGCCATCTACAGAGAGAGTCCATCTCT 1740
 Db 556 ThrGlnProSerHisLeuTyrSerAlaSerAlaIleTyrGlyGlyAsnSerSerTyrSer 575
 QY 1741 GCTTACAGCTGACGACGACGCTGCGCCACTTGGCAGACCAAGTCTTATCTGTCGACGCGG 1800
 Db 576 AlaTyrSerCysGlyClnLeuProThrCysSerAspGlnIleTyrSerValArgArg 595
 QY 1801 CAGAACCAATGACAGAGACTGACTCGCGGAGAGCTGGACATGAAGAGACCCCTTTGAA 1860
 Db 596 GlnLysProThrAspArgAlaAspSerArgArgSerThrPheGlnLysSerProPheGln 615
 QY 1861 AAGCAGTTTAAACGAGAAAGCTGCAATTTGAAATTTGAGAGACATCATGACAGAAC 1920
 Db 616 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 635
 QY 1921 AGCTCAGCGAAGACCTTGGGAGAGTGGAGAGTCACTTCTTGGGCGACATGAA 1980
 Db 636 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 655
 QY 1981 ATCATTTAGAGTCTCC 1995
 Db 656 IleIleGlnValSer 660

RESULT 5
 Q99MG6 PRELIMINARY; PRT; 677 AA.
 AC Q99MG6;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Map kinase phosphatase-M A1 isoform.
 GN DUSP16 OR 3830417M17RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
 RA Yoshikawa Y.;
 RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
 RT Activation in Macrophages.";
 RL Mol. Cell. Biol. 20:6999-7009(2001).
 RL EMBL; AF345951; AA35052.1;
 DR HSSP; Q16828; MKP.
 DR MGI; 1917936; Dusp16.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.

DR GO: GO:0005515; F:protein binding; IPI.
DR GO: GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPc; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS: PRO1764; MAPKPHPTASE.
DR SMART: SM00195; DSpc; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANES_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; Kinase.
SQ SEQUENCE 677 AA; 74550 MW; 8B6D5B7096CE2PC CRC64;

Alignment Scores:

Prod. No.:	3_15e-218	Length:	677
Score:	2954.50	Matches:	579
Percent Similarity:	92.19%	Conservative:	35
Best Local Similarity:	86.94%	Mismatches:	45
Query Match:	81.73%	Indels:	8
DB:	11	Gaps:	3

US-10-029-345A-108_COPY_538_2532 (1-1995) x Q99MG6 (1-677)

QY 1 ATGGCCCATGATGATTTGGAACCAATTGTTACTGAGAGGTTGTGGCTTGGTGA 60
Db 1 MetalahtsdluwtclleglythrghnllevalthrghluserleuValAlaleuLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTCTGCTTAATGATGACCGGCAATTTGGTAATACATACATCC 120
Db 21 SerGlyThrGluYsValleuLeuIleAspSerArgProPheValGluTyAsnThrSer 40
QY 121 CACATTTGGAGCATTAATATCACTGCTCCAAAGTTTGAAGGAGGTTGCAACAG 180
Db 41 HistleuGluAlaIleAsnIleAsnCysSerLeuMetClyAspGluLeuGlnGln 60
QY 181 GACAAAGTGTAAATATACAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGAT 240
Db 61 AspArgValleuIleThrGluLeuIleGlnHisSerIleYsIleYsValAspIleAsp 80
QY 241 TGCAGTCAGAAAGTTGATGTTACGATCAAAAGCTCCAAAGATGTTGCCCTCTCTCTCA 300
Db 81 CysAsnGlnYsValIleValIleYAspGlnSerSerGlnAspValGlySerLeuSerSer 100
QY 301 GACGTGTTTCTGCTGACTTCTGGGGTAACCTGAGAGAAGCTTCAACTGCTTCACTG 360
Db 101 AspCysPheLeuThrValleuLeuGlyLysleuGlnIleAspSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGGTGGGTTTGTGAGTTCTCTGTTGTTCCCTGGCCCTGTGAGAGAAATTC 420
Db 121 LeuAlaIleGlyIlePheAlaGlnPheSerArgCysPheProGlyLeuGlyGlySer 140
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTTGCCAATGTTGGCCAAAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 GCAATTTTCCCAATCTTTTACTTGGCTGCAGGAGATGCTCCAAAGAGGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuIleYsleuGlyCysGlnArgAspValleuAsnLysAspLeuMet 180
QY 541 CAGCAGAAATGGAGTTGTTATGTTAAATGCCAGCTATACCTGCTCCAAAGCCTGACTTT 600
Db 181 GlnGlnAsnGlyIleGlyTyValleuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGCAGATCTATTTCTGCGGTGCGCTGTGATGATGACAGCTTTGTGAGAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
QY 661 CCGGTGTTGGCAAAATCAGTATGATTTTCATTGAGAAACAAAGCCTCAATGATGTT 720

Db 221 ProThrLeuAspYsSerValAspPheIleGlnLysAlaIleYsHisSerAsnGlyCysVal 240
QY 721 CTATGTCACCTGTTTAGTGGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATG 780
Db 241 LeuIleHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIleMet 260
QY 781 AAGAGATGAGCATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 840
Db 261 LysArgMetCAspMetSerLeuAspGlnAlaTyArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCAAACTTCAATTTTCTGGGCCAACTCCCTGCACTATGAGAGAAATTAAGAC 900
Db 281 IleSerProAsnPheAsnPheMetGlyGlnLeuMetCAspYsGlySerThrIleAsnAsn 300
QY 901 CAGACTGAGCATCAGGGCCAAAGACAACTCAAGCTGCTGCACTTGGAGAAAGCCAAAT 960
Db 301 GlnThrGlyMetSerGlyProLysSerLysleuLysleuLeuHisIleuAspLysProSer 320
QY 961 GAACGTGCTGCTGCTCTCAGAGGGTGGACAGAAAGGAGAGCGCCCTCAGTCCACCC 1020
Db 321 GluProValProAlaAspSerGlnGlyTyTrpLysSerAlaLeuSerLeuSerProPro 340
QY 1021 TGTGCGACTCTGCTACTCAGAGGACGACAGACAAAGCCCGCTGATCCCGCAGCGTG 1080
Db 341 CysAlaAsnSer---ThrSerGluAlaSerGlyGlnArgLeuValHisProAla----- 357
QY 1081 CCAAGCGTCCACCGCTGACAGCCGTGCTGTTAAGAGACAGCCCGCTGTACAGCGCTC 1140
Db 358 ---SerValProSerLeuGlnProSerLeuLeuGlnAspSerThrLysLeuLysArgSerPhe 376
QY 1141 AGTGGCTGACCTGTCCCGAGACAGCGTGGAGAGACAAATAGCTCAAGCGTTCCTTC 1200
Db 377 SerGlyLeuGlnLeuSerSerGlyLysGlnAspSerThrLysLeuLysArgSerPhe 396
QY 1201 TCTCTGATATCAATACAGTTTATCATATTCAGCAGATGGCAGCATCTTACATGCTTC 1260
Db 397 SerLeuAspIleLeuSerValSerTySerAlaSerMetAlaIleSerLeuHisGlyPhe 416
QY 1261 TCTCATCAGAGATGCTTTGGAATACATACTCAAACTTCCATCTGATGGAGACCAAC 1320
Db 417 ---SerSerGlnGlnIleAlaLeuAspTyCysLysProSerAlaThrLeuAspGlyThrAsn 435
QY 1321 AAGCTATGCCAGTTCTCCCTGTTTCAGAACTATCCGAGAGAGCTCCGAAACAGTCC 1380
Db 436 LysLeuCysGlnPheSerProValGlnGluValSerGlnLeuSerProGlnThrSerPro 455
QY 1381 GATAGAGAGAAAGCCAGCATCCCAAGAACTGACAGACCGCGGCTTGCAGACAGCAG 1440
Db 456 AspLysGlnGluAlaHisIleProLysGlnProGlnProProArgProSerGlnSerGln 475
QY 1441 AGCAAGCATTTGATTCGCTCAGAACCAAGAGTGGACCGCCCAAGAGTCCCTTTTA 1500
Db 476 ValThrArgLeuHisSerValArgThrGlySerSerGlySerThrGlnArgProPhePhe 495
QY 1501 TCTCCACTGATGGAAGTGGAGCGTGGAGAGACAAATTAACAACAGCTTCTTTCCG- 1559
Db 496 SerProLeuHisArgSerGlySerValGluAspAsnIleHisThrAsnPheLeuPheGly 515
QY 1560 CTTTCCACAGCAGCAGCAGCAGCTCAAGAGTGTGCTGCGCTGAGGCTTAAAGGCTGGCA 1619
Db 516 ProPheHisGlnProAlaThrProHisGlnValCys-ArgAlaIleThrLeuLysGlyTrpHis 535
QY 1620 CTGGAATATCTGGCCCCCAGACCTTCACCTTCCCTGACCAAGACGTGATTTTGC 1679
Db 535 sSerAspIleLeuAlaProGlnSerSerAlaProSerLeuThrSerSerTrpTyArgPheAl 555
QY 1680 CACAGAGTCTTCACTTCACTGCTGCTCAGGACATCTACGAGGAGAGTGGCAGTACTC 1739
Db 555 atmGlnProSerHisLeuLysSerAlaSerAlaIleTyGlyGlyAsnSerSerTySer 575
QY 1740 TGCCTACAGCTGACAGCAGCTGCCACTTGGCGAGACCAAGTATATCTGTGCGCAGCG 1799

Db 575 rAlatYserCyGLyInLeuProThCySerSerAspGlnIleTySerValArgArgar 595
QY 1800 GCAGAGCCAGAGGACAGAGCTGACTCCGCGGAGAGCTGCGATGAGAGAGCCCTTTGA 1859
Db 595 ggInlySPromtrAspArgAlaAspSerArgArgThr-GlyMetIyArgAlaProleu 615
QY 1860 AAAGCAGTTTAAAGCCAGAGAGCTGCGCAATGATTTTGGAGAGAGCATCTGTCAGAGAA 1919
Db 615 ySerSerIeuSnInlaGlnIaIaIyStrpAnleuGlnArgAlaLeuCyArgArgT 635
QY 1920 CAGTTCACGAGAGAGCTGGGAGAGAGCTGAGTCTTACGCTTTTCCGAGAGCATGGA 1979
Db 635 hGlyProGlyArgSerTrpAlaArgTrpAlaIaIaSerProAlaSerProAlaIaIaTrpA 655
QY 1980 AATCATTGAGTCTT 1993
Db 655 rGSerSerArgSer 659
RESULT 6
Q99MG5 PRELIMINARY; PRT; 622 AA.
AC 099MG5; 301
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Map kinase phosphatase-M A2 isoform.
GN DUSP16 OR 3830417M17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]
RP SOURCE FROM N.A.
RA STRAIN=BALB/c;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.;
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages.";
RL Mol. Cell. Biol. 20:6999-7009 (2001).
DR EMBL; AF345952; AKK35053.1; -.
DR HSSP; Q16828; 1MKP.
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0000515; F:protein binding; IPI.
DR InterPro: IPR000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR001763; MAPK_phosph.
DR InterPro: IPR00387; Tyr_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHITASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE_3; 1.
DR PROSITE; PS50053; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ SEQUENCE 622 AA; 68672 MW; 6C0CA4BE909B98 CRC64;
Alignment Scores:
Pred. No.: 5,22e-192
Score: 2613.00 Length: 622
Percent Similarity: 83.93% Matches: 524
Best Local Similarity: 78.68% Conservative: 35
Query Match: 72.28% Mismatches: 45
Indels: 63
Gaps: 4
US-10-029-345A-108_COPY_538_2532 (1-1995) x Q99MG5 (1-622)
QY 1 ATGGCCATGATGATGATTGAACTCAATTTGTTACTGAGAGGTTGGTGGCTCTCTGCGAA 60

Db 1 MetAlaHleGluMetIleGlyThrGlnIleValThrGlnSerIeuValAlaLeuGln 20
QY 61 AGTGGAAACGGAAAAAGTGCTGCTAAATTGATAGCCGCGCATTTGGAAATCAATACATCC 120
Db 21 SerGlyThrGlnIleValIleuLeuIleAspSerArgProPheValGlyTrpAnthrSer 40
QY 121 CACATTTTGAAGCCATTAATATCAATGCTGCCAAGTTTGAAGAGAGGTTGCAACG 180
Db 41 HsIleLeuGlnAlaIleAsnIleAsnCySerIleuMetIySArgArgIeuGln 60
QY 181 GACAAAGTTTAATTACAGAGCTATCCAGCATTCAGCGAACAATAGTTGACATTGAT 240
Db 61 AspArgValIleuIleThrGlnIleuIleGlnHsIleSerAlaIyHsIleValAspIleAsp 80
QY 241 TGCAGTCAGAAAGTTGATTGATTACGATCAAAAGCTCCCAAGATTTGCCCTCTCTTCA 300
Db 81 CysAnGlnIySValValValIyIyAspGlnSerSerGlnAspValGlySerIeuSerSer 100
QY 301 GACTGTTTCTCACTGACTCTTGGGTTAACTGGAGAAAGCTTCACTGTTCACTG 360
Db 101 AspCySHeuThrValIleuGlnIySleuGlnIySArgSerPheAnSerValHsIleu 120
QY 361 CTTCGAGCTGGTTGCTGAGTCTCTCGTGTTCCTCGGCTCTGGAAGAAATTC 420
Db 121 LeuAla----- 122
QY 421 ACTTAGTCCCTTACCTGATTTCTAGCTTGTCTTACCTGTTGCCAATVGGGCCAAC 480
Db 122 ----- 122
QY 481 CGAATTTCCCAATTTTATCTTGCTGCGCAGAGATGTCCTCAAGAGAGCTGATA 540
Db 123 ----- 125
QY 541 CAGCAAAATGGATGTTGTTATGTTTAAATGCCAGCTATACCTGTCAAGCTGACTT 600
Db 126 GlnGlnAnGlnIyIleGlyTrValIleuAnIaSerAsnThrCySProlsProAspHe 145
QY 601 ATCCCGGAGTTCATTTCTGCGTGCGCTGCGTAATGACAGCTTTTGAGAAATTTTG 660
Db 146 IleProGlnSerIleAspIeuArgValProValAsnAspSerPheCySleuIySleu 165
QY 661 CCGTGTGTGCAAAATGATGATTTCAATTGAGAAAGCAAAAGCTTCCAAATGATGTT 720
Db 166 ProTrpIeuAspIySserValAspPheIleGlnIySalIySalIySAsnGlyCyVal 185
QY 721 CTAATGACATGTTTAACTGGAGATCTCCGCTCCGACCATGAGCTATGCTATACATG 780
Db 186 LeuIleHsCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIyIleMet 205
QY 781 AAGAGATGACATGCTTTAGATGATGATGATGATTTGAAAGAAAGAAAGCTTACT 840
Db 206 IyArgMetAspMetSerIeuSnInlaIyTrArgPheValIySglnIySArgProThr 225
QY 841 ATATCTCCCAACTTCAATTTCTGGGCCAATCTCTGAGCTATGAGAGATTAAGAC 900
Db 226 IleSerProAsnPheAsnPheMetGlyIleuMetAspIyTrGlnIyStrIleAsnAn 245
QY 901 CAGACTGAGATCAGAGGCCAAAGAGCAAACTCAAGCTGCTGCACTGAGAAAGCCAAAT 960
Db 246 GlnThrGlyMetSerGlyProIySserIySleuIySleuIySleuIySleuIySProSer 265
QY 961 GAACCTGTCCTGCTGCTGCTGAGAGGTGACAGAAAGAGAGAGCGCTCAGTCCACCC 1020
Db 266 GlnProValProAlaIaIaSerGlnIyGlyTrpIySerAlaIeuSerIeuSerProPro 285
QY 1021 TGTGCGAATCTGCTTACTCAGAGGAGAGCAAGCAAGCCGCTGATCCGCGAGCGTG 1080
Db 286 CysAlaAnSer---ThrSerGlnIaSerGlyIleArgIleValHsIleProAla----- 302
QY 1081 CCGAGCTGGCCCGAGTGGAGCGCTGCTGTTAGAGGACAGCGCGCTGACAGGCGCTC 1140

```

Db      303 ---SerValProSerLeuGlnProSerLeuGlnuAspSerProLeuValGlnAlaLeu 1321
Qy      1141 AGTGGCTGCACCTGCTCCGACAGAGCTGGAAGACAGCAATAAGCTCAAGCTTCTTC 1200
Db      322 SerGlyLeuInLeuSerSerGlnuYLeuGlnuAspSerThrYLeuYAspSerPhe 341
Qy      1201 TCTGTGATATCAAAATGATTATATTCAGCCAGCATGGAGATCTTCAACAGGCTTC 1260
Db      342 SerLeuAspLeuYSerValSerYSerAlaSerMetAlaAlaSerLeuHISGlyPhe 361
Qy      1261 TCTCATCAGAGATGCTTGGATCTAGAACCTTCACTCACTGTGATGGAGACCAAC 1320
Db      362 ---SerSerGlnuAlaLeuAspYrCyLeuYProSerAlaThrLeuAspGlyThrAsn 380
Qy      1321 AAGCTATGCAAGTTCTCCCTGTTGAGAACTATCGAGAGACATCCGAAACCAAGTCT 1380
Db      381 LeuLeuYSerGlnuPheSerProValGlnuValSerGlnuGlnSerProGluThrSerPro 400
Qy      1381 GATTAAGAGAGAGGACGACATCCCAAGAGCTGACAGCCGCGCTTCAAGACGACAG 1440
Db      401 AspYLeuGlnuAlaHISLeuProLYSGlnProGlnProAspProAspSerGlnuSerGln 420
Qy      1441 AGCAAGCATTTGATTCGTTCAGAACCCAGAGAGAGAGTGGACACCGCCAGAGTCCCTTTA 1500
Db      421 ValThrThrLeuHISerValArgThrGlySerSerGlySerThrGlnArgProPhePhe 440
Qy      1501 TCTTCATGCAATGCAAGTGGAGCGTGGAGAGACAAATTAACAACACAGCTTCTTTTCGG- 1559
Db      441 SerProLeuHISArgSerGlySerValGlnuAspAsnYrHISThrAsnPheLeuPheGly 460
Qy      1560 CTTTTCACAGCAGAGAGACCTCCAGAGAGTCTGCGTGGGCTTGAAGGCTTGGCA 1619
Db      461 ProPheHISGlnProAlaThrProHISGlnValCys-ArgAlaThrLeuYSerGlyThrPhe 480
Qy      1620 CTCGATATCTTGGCCCCCGACAGCTTACCCCTTCCCTGACAGAGAGTGTATTTTGC 1679
Db      480 SerSerPheLeuAlaLeuAlaProGlnSerSerAlaProSerLeuThrSerSerThrPyrPheAl 500
Qy      1680 CACAGAGCTCTCACATTTACTCTGCTGCTCAGCCATTCAGAGAGAGAGTCCAGTTACTC 1739
Db      500 AtHISGlnProSerHISLeuYrSerAlaSerAlaLeuYrGlyYAspSerSerYrSe 520
Qy      1740 TGGCTACAGTGCAGCGAGCTGCGCACTTGCGAGAGACAGTATCTGTTGGGACGGC 1799
Db      520 rAlaYrSerCyGlyGlnuLeuProHISCySerAspGlnuLeuYrSerValArgPheArg 540
Qy      1800 GCAGAAGCCAGTGCAGAGCTGACTGCGCGGAGTGGCATGAAGAGAGCCCTTTGA 1859
Db      540 GglnuYrProThrAspArgAlaAspSerAlaArgThr-GlyMetLYAspArgAlaProLeuL 560
Qy      1860 AAAGCATTTAAACGAGAAAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCCAGAGAA 1919
Db      560 YSerSerLeuAsnAlaGlnuAlaAlaLeuYrTrpAsnLeuGlnuArgAlaLeuCYAspArgTr 580
Qy      1920 CAGGTACCGGAGAGAGCTGGGAAAGTGGGACATGATCTTTCCGGGACAGATGA 1979
Db      580 HTrGlyProGlyArgSerTrpAlaArgTrpAlaAlaSerProAlaSerProAlaAlaTrp 600
Qy      1980 AATCATTTAGAGTCT 1993
Db      600 rGSerSerArgSer 604

```

```

OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxId=10990;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BALB/c;
RA      Matsuguchi T., Musikacharoen T., Johnson T.R., Krafe A.S.,
RT      Yoshikai Y.;
RT      "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT      Activation in Macrophages."
RL      Mol. Cell. Biol. 20:6999-7009(2001).
DR      EMBL; AF345953; AAK35054.1; -.
DR      HSSP; Q16828; IMKP.
DR      MG1;1917936; Dusp16.
DR      GO; GO:0005737; Cytoplasm; IDA.
DR      GO; GO:0005634; Cytosol; IDA.
DR      GO; GO:0005515; F-protein binding; IPT.
DR      GO; GO:0000188; P-activation of MAPK; IDA.
DR      InterPro; IPR000340; DS phosphatase.
DR      InterPro; IPR008343; MAPK phosph.
DR      InterPro; IPR001763; Rhodanese-like.
DR      InterPro; IPR000387; TYR_phosphatase.
DR      Pfam; PF00782; DSPC; 1.
DR      Pfam; PF00581; Rhodanese; 1.
DR      SMART; SM00195; DSPC; 1.
DR      SMART; SM00195; DSPC; 1.
DR      PROSITE; PS50206; RHODANASE_3; 1.
DR      PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR      PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR      PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM      Hydrolase; Kinase.
SQ      SEQUENCE 355 AA; 39502 MW; D52C29AE215CA285 CRC64;

Alignment Scores:
Pred. No.: 3,58e-103 Length: 355
Score: 1455.50 Matches: 314
Percent Similarity: 49.02% Conservative: 12
Best Local Similarity: 47.22% Identities: 11
Query Match: 40.26% Mismatches: 328
DB: 11 Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) x Q9AG16 (1-355)
Qy      1 ATGGCCCATGAGATGATTTGAATCAATTTTACTGAGAGGTTGGTGGCTGCTGGAA 60
Db      1 MetAlaHISGlnuLeuLeuGlnuGlnuLeuValThrGlnuSerLeuValAlaLeuLeuGln 20
Qy      61 AGTGAACGGAAAAAGTGTGCTAATTTGATAGCGGCAATTTGGTAATCAATATCATCC 120
Db      21 SerGlyThrGlnuYrValLeuLeuLeuLeuLeuAspSerArgProPheValGlnuYrAsnThrSer 40
Qy      121 CACATTTTGAAGCCATTAATATCAATGCTGCCAAGCTTATGAAGAGAGGATTTGAT 180
Db      41 HISLeuGlnuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeu 60
Qy      181 GACAAAGTGTATTAACAAGCTCAAGCTCAATCCAGCATTCAGCCGAAACATAGAGTTGAT 240
Db      61 AspArgValLeuLeuLeuThrGlnuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Qy      241 TGCAGTCAGAGGTTAGTTTACATGATCAAGCAAGCTCCCAAGATGTTGCTCTCTTCA 300
Db      81 CyAsnGlnuYrValValValValValValValValValValValValValValValValValVal 100
Qy      301 GACTGTTTTCTCACTGACTTCTGGGTAACCTGAGAGAGCTTCAACTGTTCACCTG 360
Db      101 AspCyPheLeuThrValLeuLeuGlnuYrLeuGlnuArgSerPheAsnSerValHISLeu 120
Qy      361 CTTCGAGGTGGGTTGCTGATGTTCTGTTGTTTCCGCGCTCTGTAAGGAAATCC 420
Db      121 LeuAlaGlyGlyPheAlaGlnuPheSerArgCyPheProGlyLeuYrGlnuYrGlnuYrSer 140
Qy      421 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGTGTTACTGTTGCAATTTGGGCAACC 480

```



```

Db      141 ThrLeuValProthrCysIleSerGlnProCysLeuProValAlaIleuIleGlyProThr 160
QY      481 CGAATTCCTCCCAATTTTATCTTGGCTGCCAGCAGATGTCTCCCAACGAGCTGATA 540
Db      161 ArgIleuProAsnLeuIleuGlyCysGlnArgSerValLeuAsnLeuSapLeuMet 180
QY      541 CAGCGAATGGATGGTTATGTGTAAATGCCAGCTATACCTGTCCAAAGCTGACTT 600
Db      181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProIleProAspHe 200
QY      601 ATCCCGAGTCTCATTTCCCTGCTGCTGTGTGATGATGACGCTTTGTGAGAAATTTTG 660
Db      201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIleLeu 220
QY      661 CCGTGGTTGACAAATACAGATGATTGATGAGAAAGCAAAAGCTCCCAATGATGTGT 720
Db      221 ProThrLeuAspIleSerValAspPheIleGlnValAlaValAlaSerAsnGlyCysVal 240
QY      721 CTATGACACTGTTAGTGGATCTCCGCTCCGACCATGCTATGCTGCTACATCATG 780
Db      241 LeuIleHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY      781 AAGGAGTGGACATGTCTTTAGATGAGCTTACAGATTTGTGAAAGAAAGAAAGCTACT 840
Db      261 LysArgMetAspMetSerLeuAspGlnAlaTyr----- 271
QY      841 ATATCTCCAAACTTCAATTTTCTGGCCCACTCTGACATATGAGAAAGATTAGAAC 900
Db      271 ----- 271
QY      901 CAGACTGAGCATCAGGGCCAAAGACAAACTCAGCTGTGACCTGGAGAAAGCCAAAT 960
Db      271 ----- 271
QY      961 GAACCTGTCCCTGTGTCTCAGAGGCTGACAGAAAGCGAGAGCGCCCTCAGTCCACC 1020
Db      271 ----- 271
QY      1021 TGTGCCACTCTGCTACTCCTAGAGGACAGAGCAAAAGCCGTGATCCGCCAGCGTG 1080
Db      271 ----- 271
QY      1081 CCCAGCGTCCAGCGTGCAGCGCGTGTGTAGAGAGACGCCCGTGTGACAGCGCTC 1140
Db      271 ----- 271
QY      1141 AGTGCGTGCACCTGTCCGACAGCGCTGGAAGACGATATAGCTCAAGCTTCTTC 1200
Db      271 ----- 271
QY      1201 TCTCTGATATCAATCAGTTTCATATTCAGCCAGCATGCGACATCCTTACATGCTTC 1260
Db      271 ----- 271
QY      1261 TCCTCATCAGAAAGATGCTTTGGAATATCAAAACCTTCACTACTGTGATGGAGCAAC 1320
Db      271 ----- 271
QY      1321 AAGCTATGCCAGTCTCCCTGTTCCAGAACTATCGAGAGACACTCCCGAAACAGTCTC 1380
Db      271 ----- 271
QY      1381 GATAGAGAGAAAGCAGCATCCCAAGACCTGACACCGCCAGGCTTCAGACAGCCAG 1440
Db      271 ----- 271
QY      1441 AGCAAGCATTTGCAATTGCGTCAAGAACGACAGAGATGAGACCGGCCAAGAGTCCCTTTTA 1500
Db      271 ----- 271
QY      1501 TCTCACTGATGAGAGTGGAGCGTGGAGACAAATTACACACCAAGCTCTTTTGGC 1560
Db      271 ----- 271

```

```

QY      1561 CTTTCCACGACGACGACAGCACTCAGAAAGTCTGTGCTGGCTTAAGGGCTGGCAC 1620
Db      271 ----- 271
QY      1621 TCGAATATCTTGAGCCGCCACGACTCTACCCCTTCCCTGACGACAGCTGATTTTGGC 1680
Db      271 ----- 271
QY      1681 ACAGAGTCTCACAATTCTACTGTGCTGACCATCTTACGAGAGCAGTGCCAGTTACTCT 1740
Db      271 ----- 271
QY      1741 GCTTACAGCTCAGCAGCTGCCCACTTGGGAGACCAAGTCTATTTCTGCGCCAGCGG 1800
Db      272 ----- 272
QY      1801 CAGAACCAAGTGCAGAGGCTGATCGCGCGGAGCTGGCATGAAGAGCCCTTTGAA 1860
Db      274 GlnIysProThrAspAlaAspSerArgArgThr-GlyMetLysArgAlaProLeuLys 1920
QY      1861 AAGCATTTTAAAGCAGAAAGCTGCCAATGAAATTTGGAGAGCATCATGTCAGAGAC 1920
Db      293 SSerSerLeuAsnAlaGlnAlaAlaIleTyrAsnLeuGlnArgAlaLeuCysArgArgThr 313
QY      1921 AGGTACCGGAGAGAGCTGGGAAAGTGGGCACTGCTAGCTTTTGGCGAGCATGAA 1980
Db      313 rGlyProGlyArgSerTyrAlaArgTyrAlaAlaSerProAlaSerProAlaIleTyrPar 333
QY      1981 ATCATTAGAGTCT 1993
Db      333 gSerSerArgSer 337

```

RESULT 8

```

ID      086SS8      PRELIMINARY;      PRT;      625 AA.
AC      086SS8;
DT      01-JUN-2003 (TREMBLrel. 24, Created)
DT      01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT      01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE      Similar to dual specificity phosphatase 8.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      Tissue=Brain, and Astrocytoma;
RA      Strausberg R.;
RL      Submitted (JAN-2003) to the EMBL/Genbank/DBJ databases.
DR      EMBL; BC045110; AA045110.1; -.
DR      GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR      GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR      InterPro; IPR000340; DS phosphatase.
DR      InterPro; IPR008343; MAPK phosph.
DR      InterPro; IPR002965; P-rich extensn.
DR      InterPro; IPR001763; Rhodanese-like.
DR      InterPro; IPR000387; TYR_phosphatase.
DR      Pfam; PF00782; DSPC; 1.
DR      PRINTS; PRO1764; MAPKPHPTASE.
DR      PRINTS; PRO1217; PRICHEXTENSN.
DR      SMART; SM00195; DSPC; 1.
DR      SMART; SM00450; RHOD; 1.
DR      PROSITE; PS50206; RHODANESE_3; 1.
DR      PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR      PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR      PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ      SEQUENCE 625 AA; 65826 MW; C7C80840B724PFC CRC64;

```

Alignment Scores:

```

Ptd. No.:      3,13e-93      Length:      625
Score:         1327.00      Matches:     314
Percent Similarity: 58.54%      Conservative: 94

```


Best Local Similarity: 45.05% Mismatches: 185
Query Match: 36.71% Indels: 104
DB: 4 Gaps: 21
US-10-029-345a-108_copy_538_2532 (1-1995) x Q86S88 (1-625)

QY 1 ATGGCCCATGATGATGGAACATTAATGTT---ACTGAGAGTTGGTGGCTGCTG 57
DB 1 MetAlaGlyAspArgLeuProAlaGlyValMetAspAlaIlyblybLeuAlaSerLeu 20
QY 58 GAAAGTGAAAGGAAAAAGTGGCTTAATGATAGCCGACATTTGTAATCAATACA 117
DB 21 ArgGlyGlyProGlyGlyProLeuValIleAspSerArgSerPheValGlyIlyrAsnSer 40
QY 118 TCCCACTTTTGGAGCCATTAATATGATCTCCCAAGCTTATGAAAGCGAAGTTGCA 177
DB 41 TrpIleValLeuSerValAsnIleCysCysSerIlyblybValIlybAspGlyLeuGln 60
QY 178 CAGGACAAAGTGAATTAATACAGAGCTCATCAGATTCAGGAAACATTAATGACATT 237
DB 61 GlnGlyLysValThrIleAlaGlyLeuIleGlnProAlaIleAspSerGlnValAla 80
QY 238 GATTGACGACAGAGTTGATGATTAAGATCAAGCTCCCAAGATGTTGCTCTCTCT 297
DB 81 ThrGluProGlnAspValValIlyrAspGlnSerThrArgAspAlaSerValLeuAla 100
QY 298 TCAAGACTGTTTCTCACTGATCTTGGGTAACGTGAGAAAGACTTCAACTGTTTCA 357
DB 101 AlaAspSerPheLeuSerIleLeuLeuSerIlyblybCysPheAspSerValAla 120
QY 358 CTGCTGACAGAGGAGTTGCTGATGTTCTCGTTGTTCCCTGGCCCTGTAAGGAAA 417
DB 121 IleLeuThrGlyGlyIleAlaThrPheSerSerCysPheProGlyLeuGlyGlyLys 140
QY 418 ---TCCACTGTAGTCCCTACCTGATTTCTCAGGCTTCTTACCTGTTGCAACATTGG 474
DB 141 ProAlaIleLeuLeuProMetSerLeuSerGlnProCysLeuProValProSerValGly 160
QY 475 CCAACCGAATTTCTCCCAATCTTTATCTTGGCTGCCAGAGATGTTCTCAACAGAG 534
DB 161 LeuThrArgIleLeuProIleLeuIlyrLeuGlySerIlnIlybAspValIleAsnIlybAsp 180
QY 535 CTGATACAGAGATGGGATGTTGTTAAATGCGCAACTAACCTCCCAAGGCT 594
DB 181 LeuMetThrGlnAsnGlyIleSerIlyrValLeuAsnAlaSerAsnSerCysProLysPro 200
QY 595 GACTTTATCCCGAGTCTCATTTCTGCTGCTGCTGCTGATGACAGCTTTTGAGAAA 654
DB 201 AspPheIleCysGlySerArgPheMetArgValProIleAsnAspAsnIlyrCysGlyLys 220
QY 655 ATTTTGGCGTGTGGAACAATCAGATGATTTCAATGAGAAAGCAAGCTCCCAATGA 714
DB 221 LeuLeuProIlybLeuAspIlybSerIleGlnPheIleAspIlybAlaIlybLeuSerSerCys 240
QY 715 TGGTCTAGTGAAGCTTTAGAGTCTCCGCTCCGCGCAACGATCGTATCGGCTAC 774
DB 241 GlnValIleValIleCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIlyr 260
QY 775 ATCATGAGAGATGACATGCTTTTATGATGAGCTTACAGATTTTGAAGAAAAGA 834
DB 261 IleMetIlyrThrMetGlyMetSerSerAspAspAlaIlyrArgPheValIlybAspArgArg 280
QY 835 CCTACTATATCTCCAACTTCAATTTTCTGGGCAACTCTGACATATGAGAAAGATT 894
DB 281 ProSerIleSerProAsnPheAsnPheLeuGlyIleGlnLeuGlnIlyrGlyArgSerLeu 300
QY 895 AAGAACAGACTGAGAGATCAGGGGCAAAAGCAAACTCAAGCTGCGACCTGAGAGAG 954
DB 301 LysLeuLeuAlaIleAlaLeuGlnIlybAspProGlyIlyr-----Thr 312
QY 955 CCAAAAT-----GAACCTGTCTCTGTCTCAGAGGATGAGACAGAAAGAGAGAGC 1005
DB 313 ProSerGlyThrProGluProProProSerProAlaIleGlyAlaProLeuProArgLeu 332

QY 1006 CCCCTCAGTCCACCCCTGTGCGGACTGTGCTAC-----TCAGAGGACAGAGCAAGG 1059
DB 333 Pro-----ProThrThrSerGlnSerAlaAlaThrGlyAsnAlaIleAlaGlyGly 350
QY 1060 CCCGTGCATCCCGCCGCGGTGCGCCAGCTGCGCAGCGTGCAGCTGCTGTTAGAGAC 1119
DB 351 GlyLeuSerAlaGlyGlyGlyProProAlaProProThrProProAla-----Thr 367
QY 1120 AGCCCGCTGATCAGCGCTCACTGGCTGACCTGTCCGACAGACAGGCTGGAAGACAGC 1179
DB 368 SerAlaLeuGlnGlnIlyLeuArgGlyLeuAsnIleSerSerAspArgLeuGlnAspThr 387
QY 1180 AATTAAGCTCAAGGCTCTCTCTCTGATATCAATCAATGTTTCATATTACGACGATG 1239
DB 388 AsnArgLeuIlybArgSerPheSerLeuAspIleLysSerAla----- 401
QY 1240 GCAGACTCCCTTACATGCTTCTCTCATCAGAAAGATGTTGAAATACATAAAGCTTCC 1299
DB 402 -----TyrAlaProSer 405
QY 1300 ACTACTCTGATGAGCAAC-----AACTATGCCAGTTT 1335
DB 406 ArgArgProAspGlyProGlyProProAspProGlyGlyAlaProLysLeuGlybLeu 425
QY 1336 ---TCCCTT-----GTTCAAGAACTATGCGAGACAGCTCCGAAACAGTCTGAT 1383
DB 426 AspSerProSerGlyAlaIleAlaLeuGlyLeuSerSerProSerProAsp---SerProAsp 444
QY 1384 AAGGAGAAAGCCAGCATCCCAAGAACTGACAGCCGACGCTTACAGACCCAGAGC 1443
DB 445 AlaAlaProGlyAlaArgProArgProArgArgProArgPro----- 459
QY 1444 AAGCGATTGCAATTCGTCAGAACACAGACAGATGCGACCGCCAGAGTCTTTATCT 1503
DB 460 -----ProAlaGlySerProAlaArg---Ser 467
QY 1504 CCACTGCATCGAAGTGGAGCGGTGAGAGCAATTACCAACAGCTTCTTTTCCG--- 1560
DB 468 ProAlaIleSerLeuIlyr-----LeuAsnPheGlyAsp 478
QY 1561 ---CTTCCACAGCCAGACAGACACTCAGCAATCTGCTGCGCTTAAAGGCG 1614
DB 479 AlaAlaArgGlnThrProArgIleGlyLeuSerAlaLeuSerAlaProGlyLeuProGly 438
QY 1615 -----TGGCACTCGAATATCTTGGCCCCCAGACCTCT 1647
DB 499 ProGlyGlnProAlaGlyProGlyAlaIlyrAlaProProLeuAspSerPro-----Gly 516
QY 1648 ACCCTTCCCTGACACAGACAGCTGATTTTCCACAGAGTCTTCAACTTACTCTGCC 1707
DB 517 ThrProSerProAspGlyProIlyrCysPheSerProGly----- 529
QY 1708 TCAGCATCTACAGGAGAGTGGCACTGCTGCTTACAGCTGACAGCCAGCTGCCACT 1767
DB 530 GlyAlaGlnGlyAlaIlyGlyValLeuPheAlaProPheGlyArgGlyAlaProGly 549
QY 1768 TGGGACCAAGTCAATTTCTGTGCGAGCGGAGCAAGCCAAAGTGAAGCTGACTG 1827
DB 550 ProGlyGlyGlySerAspLeuArgArgGlyAlaAlaArgGlyAlaIlyrProArgAspAla 569
QY 1828 CGCGGAGCTGCGATGAAGAGAGCCCTTTGAAAGCACTTAAAGCGAAGCTGCCAA 1887
DB 570 ArgThrGlyIlyrProGluGluProAlaProGluThrGlnPheIlybArgArgSerCysGln 589
QY 1888 ATGAATTTGAGAGAGCATCATGTCAGAGAAACAGTCAAGC-----GAAGAGCTGGGAGAA 1944
DB 590 MetGluPheGluGluIlyr---MetValGlyGlyArgAlaArgGlyGlyGlyLeuAlaAla 608
QY 1945 GTGGGCACTGACTTACGTTTTCGGGACAGAGAAATCAATGAGGTTCTC 1995
DB 609 LeuGlyLysGlnAlaSerPheSerGlySerValGlyValIleGlyValSer 625


```

QY 1330 CAGTTCCTCCCTGTCAGAACTATCGAGACAGCTCCGAAACAGTCTGATGAAGAG 1389
DB 438 -----ProSerProSerProSerPro 445
QY 1390 GAAGCCAGCATCCCAAG-----AGCTGCAGACCCGACGCTTCAGACAGCAG 1440
DB 446 Asp---SerValProGluCyArgProArgProArgArgArgProProAlaSerSer 464
QY 1441 AGCAAGCGA-----TTGCATTCGGGTGAGAACCAAGCAGAGTGGCACCCTCCAGAGTCC 1494
DB 465 ProAlaArgSerProAlaArgLeuGlyLeuAlaSerProAlaArgThrAlaArgGln--- 483
QY 1495 CTTTATCTCAGCAGCTCGAAGTGGAGGAGGTGGAGGACAATTACACAGCAGTTCCTT 1554
DB 484 -----ThrProArgHis----- 487
QY 1555 TTGGGCTTTTCACACAGCAGCAGACCTTCAGAGTCTGCGCTG----- 1602
DB 488 ---GlyLeuSerAla-----LeuSerAlaProGlyLeuProGlyProGly 501
QY 1603 -----GGCTTAAAGGCTGGCACTCGGATATCTGGCCCCCAGACCTTACCCCT 1653
DB 502 GlnProAlaGlyProGlyGlyTyrPvalProProLeuAlaSerPro-----GlyThrPro 519
QY 1654 TCCCTGACCAAGCAGCTGTATTTTGCCACAGAGTCTCACAC-----TTCTAC 1701
DB 520 SerProAlaGlyProTyrCyPheSerProGlnGlyAlaGlnGlyProGlyAlaValPhe 539
QY 1702 TCTGCC-----TCAGCCATCTACGGA 1722
DB 540 SerAlaAlaPheGlyArgValSerAlaGlyAlaProGlyProGlyAlaSerSerSerGly 559
QY 1723 GGC----- 1725
DB 560 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 579
QY 1726 -----AGTGCAGTACTCTGCTACAGCTGCAGCCAGCGCCACTTGGGAGAC 1776
DB 580 SerAlaSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 599
QY 1777 CAACTATTCTGTGCGCAGCGCAGAGCAAGTGCAGAGCTGACGCGCGAGC 1836
DB 600 SerSerSerAlaPheArgArg-----AspValArgThrGly 612
QY 1837 TGGCAGTGAAGAGCCCTTTGAAAGCAGTTTAAAGCCAGAGCTGCCAAATGGAATTT 1896
DB 613 TrpProGlnLupProAlaAlaAspAlaGlnPheGlyArgSerCyGlnMetGlnPhe 632
QY 1897 GGAGAGAGCATCATGTGCAGAGAACAGTCAAGG---GAAGAGCTGGGGAAGTGGCGCAGT 1953
DB 633 GlnGlnGly---MetValGlnGlyArgAlaArgGlyGlnGlnGlnAlaAlaLeuGlyLys 651
QY 1954 CAGTCTAGCTTTTCGCGCAGCAGCAGAAATCATTTGAGGCTCC 1995
DB 652 GlnThrSerPheSerGlySerValGlnValIleGlnValSer 665

```

```

RA Yoshikai Y.;
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages."
RL Mol. Cell. Biol. 20:6999-7009 (2001).
DR EMBL; AF345954; AAK35055.1; -.
DR HSSP; Q16828; MKP.
DR MGD; MGI:1917936; Duple16.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005634; Cytosol; IDA.
DR GO; GO:000515; F:protein binding; IPT.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ SEQUENCE 300 AA; 33624 MW; 4C61846ACDF0F456 CRC64;

Alignment Scores:
Pred. No. 5,79e-77 Length: 300
Score: 1114.00 Matches: 259
Percent Similarity: 40.75% Conservative: 12
Best Local Similarity: 38.95% Mismatches: 11
Query Match: 30.82% Indels: 383
Gaps: 2

US-10-029-345A-108_COPY_538_2532 (1-1995) x Q9AG15 (1-300)
QY 1 ATGGCCCATGAGATGATGGAATCAATTTGTTACTGAGAGGTGGTGGCTGTGGA 60
DB 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnSerLeuValAlaLeuGln 20
QY 61 AGTGAACGGAAGAAATGCTGCTAATTTGATAGCCGCGCAATTTGTGGAATACATATCC 120
DB 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheAlaGlnTyrAlaThrSer 40
QY 121 CACATTTGGAACCATTAATATCAATGCTCCCAAGCTTATGAGCGAAGTTGCAACAG 180
DB 41 HisIleLeuGlnAlaIleAlaSerIleAsnCySerLysLeuMetLysArgArgLeuGln 60
QY 181 GACAAGGTGTAATTCAGAGCTCATCCAGCATTCAGCGAACAATGAAGTTGACATTTGAT 240
DB 61 AspArgValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAlaPheAsp 80
QY 241 TGAAGTCAGAGGTTTGAATTCAGATCAAGCTCCCAAGATGTTCCTCTCTCTCA 300
DB 81 CysAlaGlnLysValValValTyrAspGlnSerSerGlnAspValGlySerLeuSerSer 100
QY 301 GACTGTTTCTCAGTACTCTTGGGTAACTGGAAGAGCTTCACTCTGTTCACTG 360
DB 101 AspCyPheLeuThrValLeuLeuGlyLysLeuGlnArgSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGAGGTGGTTCGATGTTCTCGTTGTTTCCCTGCGCTCTGTGAAGAAATCC 420
DB 121 LeuAla----- 122
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCTTGCTTACTGTTGCCAATTTGGGCAACC 480
DB 122 ----- 122
QY 481 CGAATCTTCCCAATCTTATCTTGCGTCCAGCGAGATGTCTCAACAGAGCTGATN 540
DB 123 -----AspLeuMet 125

```


Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnSerLeuValAlaLeuGlu 20
QY 61 AGTGAACCGAAAAAGCTGCTTAATGATAGCCGGCATTGTGGAAATACATACATCC 120
Db 21 SerGlyThrGlnValLeuLeuIleAspSerArgProPheValGlyIleAsnThrSer 40
QY 121 CACATTTGGAGCATTAAATATCACTGCTCCCAAGTTTGAAGGAGGTTGGCAACAG 180
Db 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerLeuMetCysArgArgLeuGlnGln 60
QY 181 GACAAAGCTGTAATTAACAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
Db 61 AspIleValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATGATTTACGATCAAGCTCCCAAGATGTGGCTCTCTCTTCA 300
Db 81 CysAsnGlnIleValValIleValIleValIleValIleValIleValIleValIleVal 100
QY 301 GACTGTTTCTCACTGACTCTTGGGTTAACTGGAGAGAGCTTCAACTGTTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCGACAGTGGGTTTGTGCTGATCTCTGTTGTTTCCCTGCTGTGAGAGAAATCC 420
Db 121 LeuAlaGlyLysPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlnGlyLysSer 140
QY 421 ACTCTAGTCCCTTACCTGATTTCTGAGCTTGTGCTTGTGCTTGTGCAACATGGGCCAAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCACTTTATCTTGGCTGCGAGAGAGATGCTCCCAAGAG 531
Db 161 ArgIleLeuProAsnLeuIleIleGlyCysGlnArgAspValLeuAsnLys 177
RESULT 12
Q96GS2 PRELIMINARY; PRT; 143 AA.
ID 096GS2
AC 096GS2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Truncated MAPK phosphatase 7.
GN MKP7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Montpetit A., Boily G., Simmet D.;
RT "A detailed transcriptional map of the chromosome 12p12 tumor
suppressor locus."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY038927; AAK69770.1; -
DR GO; GO:0017017; F-MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P-protein amino acid dephosphorylation; IEA.
DR InterPro; IPR008343; MAPK_Phospn.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00450; RHOD.
DR PROSITE; PSS0206; RHODANES_3; 1.
SQ SEQUENCE 143 AA; 16077 MW; 5213A213AA7E5974 CRC64;

Alignment Scores:

Pred. No.: 1,2e-41 Length: 143
Score: 653.00 Matches: 140
Percent Similarity: 70.35% Conservative: 0
Best Local Similarity: 70.35% Mismatches: 3
Query Match: 18.06% Indels: 56
DB: 4 Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) X Q96GS2 (1-143)

QY 1 ATGGCCATGAGATGATGGAACTCAATTTGTTACTGAGAGGTTGGGCTGTGCGAA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnSerLeuValAlaLeuGlu 20
QY 61 AGTGAACCGAAAAAGCTGCTTAATGATAGCCGGCATTGTGGAAATACATACATCC 120
Db 21 SerGlyThrGlnValLeuLeuIleAspSerArgProPheValGlyIleAsnThrSer 40
QY 121 CACATTTGGAGCATTAAATATCACTGCTCCCAAGTTTGAAGGAGGTTGGCAACAG 180
Db 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerLeuMetCysArgArgLeuGlnGln 60
QY 181 GACAAAGCTGTAATTAACAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
Db 61 AspIleValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATGATTTACGATCAAGCTCCCAAGATGTGGCTCTCTCTTCA 300
Db 81 CysSerGlnIleValValIleValIleValIleValIleValIleValIleValIleVal 100
QY 301 GACTGTTTCTCACTGACTCTTGGGTTAACTGGAGAGAGCTTCAACTGTTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCGACAGTGGGTTTGTGCTGATCTCTGTTGTTTCCCTGCTGTGAGAGAAATCC 420
Db 121 LeuAlaGly--Ala----- 124
QY 421 ACTCTAGTCCCTTACCTGATTTCTGAGCTTGTGCTTGTGCTTGTGCAACATGGGCCAAC 480
Db 124 ----- 124
QY 481 CGAATCTTCCCACTTTATCTTGGCTGCGAGAGAGATGCTCCCAAGAGCTGATA 540
Db 125 -----AspA 126
QY 541 CAGCAAGATGGATGTTGTTATGTTAATGCCAGCTATACCTGCCAAAGCC 593
Db 126 IaAlaGlnIlePheIlePheLeuValIleValIleValIleValIleValIleValIleVal 143
RESULT 13
Q8ST19 PRELIMINARY; PRT; 657 AA.
ID Q8ST19
AC Q8ST19;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F08B1.1a.
GN F08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans a platform for
investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Chisoe S.;
RT "The sequence of C. elegans cosmid F08B1.1";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U23178; AAK68299.1; -.
 DR HSSP: Q16828; IMKP.
 DR WormPep: F08B1.1a; CE27918.
 DR GO: GO:0016787, F-hydrolase activity, IEA.
 DR GO: GO:0017017, F-MAP kinase phosphatase activity, IEA.
 DR GO: GO:0006470, P-protein amino acid dephosphorylation, IEA.
 DR InterPro: IPR000340, DS_phosphatase.
 DR InterPro: IPR008343, MAPK_phosph.
 DR InterPro: IPR000387, TYR_phosphatase.
 DR InterPro: IPR001763, Rhodanese-like.
 DR Pfam: PF00782, DSPc, 1.
 DR PRINTS: PR01764, MAPKPHNTASE.
 DR SMART: SM00195, DSPc, 1.
 DR SMART: SM00450, RHOD, 1.
 DR PROSITE: PS50206, RHODANASE_3, 1.
 DR PROSITE: PS00383, TYR_PHOSPHATASE_1, 1.
 DR PROSITE: PS50056, TYR_PHOSPHATASE_2, 1.
 DR PROSITE: PS50054, TYR_PHOSPHATASE_DUAL, 1.
 DR Hypothetical protein: Hydroxalase
 Q0 SEQUENCE 657 AA: 71002 MW: 106040 Da
 ...

Alignment Scores:

Pred. No.:	1,428-39	Length:	657
Score:	282.50	Matches:	206
Percent Similarity:	46.42%	Conservative:	99
Best Local Similarity:	31.35%	Mismatches:	212
Query Match:	17.39%	Indels:	140
DB:	5	Gaps:	25

88-10-025-345A-108_COPY_538_2532 (1-1995) X Q8ST19 (1-657)

```

OY      28  ATGTGTAAGTGAAGGTTGCTGCTGCTGTGGAAAAGTGGAAACGGAATAAGTCCGTAATT  87
Db      9  TlSeSrThCySeGlyLeuAlaLeuIleArgGluIaIaProAspThrLeuValVal
OY      88  GATAGCCGGCCCAATTTGTGGAAATPACAATCAATCCCACTTTTGGAAAGCCATTAATCAAC  14
Db      29  AspCyArGArgGlyPheThrGluIuArgAsnIuSerHisValArgHisSerMetAsnIaPhe  48
OY      148  TGCTCCAGCTTATTAAGCGAAGGTTGCAACAGAGCAAAAGT-----TTATT  19
Db      49  PheSerLysLeuIleArgArgLeuPheGluAsnLysLeuAspAsnCySeuIle  68
OY      196  ACAGAGCTCATCAGCATTACGCG-----AAACATPAGGTTGACATTGAT  240
Db      69  HisGluLeuMetSerCySeSerGlyCyStrIlyMetAspGluLysLeuAspLeu---  87
OY      241  TGCAGCTCAGAAAGTTGAGTTAGTACATCA-----AGCTCCCAAGAT  282
Db      88  -----ValLeuIyPAlaGluGluAspLysProArgGlyAsnLysArgArg  102
OY      283  GTTGCCCTCTCC-----TCTTACAGACTGTTTCTCAGCTACTTCTGCGTAA  330
Db      103  IleIaSerCyAsnAlaProGluSerThrAlaLysIleMetArgValLeuArgGluArg  122
OY      331  CTGAGAAGAGC-----TTCACTGTGTTCAACGCTGTGAGGTGGTGTGCTGAGTTC  384
Db      123  LeuGluAspThrAspLysPheArgSerValMetValLeuGluGlyGlyPheLysGlnPhe  142
OY      385  TCTCGTGTTCCTCGGCTCTGTGAAGAAATCC-----ACTTAAGCCCTACTCTGC  438
Db      143  AlaIleIuIuArgProGluLeuCySeGluSerGluGlyMetThrArgLeuProGluSer  162
OY      439  ATTTTCAGGCTTGCTTA-----CCTGTTCCAACATTTGGGCCAACCGAATTTCTCC  492
Db      163  LeuSerGlnProCySeuSerGlnProThrGlyAsp---GlyIleThrLeuIleThrPro  181
OY      493  AATCTTATCTTGGCTGCCAGCAGATGTCCTCAACAAGAGCGATATACAGACGATGG  552
Db      182  AsnIleIyLeuGluSerGlnIleAspSerLeuAspGluIuThrMetLeuAspAlaLeuAsp  201
OY      553  ATTGGTATGTGTAAATGACAGCTATACCTGTCCAAAGCTGACATTATCCCGAG---  609

```

Db 202 ILeserValIleIleAsnLeuSerMetThrCysProIysSerValCysIleLeuGluAsp 221
 QY 610 TCTCATTTCCGTGGTGGCCGTGTGATGACAGCTTTTGTGAGAAATTTTCCGTGGTTG 665
 Db 222 LysAsnPhMetArgIleProValAsnAspSerIyrGlnGluIlybLeuSerProIyrPhe 241
 QY 670 GACAAATCAGTGAATTTATGTGAGAAACCAAAAGCTTCAATGGATGTGTTTCAATGCAC 729
 Db 242 ProMetAlaItycIluPheIleuIlybCysArgIyrGlaIyIlybIysCysLeuIleHis 261
 QY 730 TGTTTAGCTGGGATCTCCCGCTCCGCACACATCCCTATCGCTTACATCATGAGAGAGAG 789
 Db 262 CysLeuAlaGlyIleSerArgSerProThrIleuAlaIleSerIyrIleMetArgIyrMet 281
 QY 790 GACATGCTTTTGAATGAAGCTTACAGATTTGTGAAAGAAAAAGCTTACTATATCTCA 849
 Db 282 LysMetGlySerAspAspAlaIyrArgIyrValIyGluIlyArgIyrProSerIleSerPro 301
 QY 850 AACTTCATTTTCTTGGCCAACTCTCGGACTATGAGAAAGATGTAAAGACACAGACTGGA 909
 Db 302 AsnPheAsnPhMetGlyGlnLeuLeuGluIyrGluAsnValLeu----- 316
 QY 910 GCATCAGGCCCAAGACAACTCAAGCTGCGACCTGGAGAACCCAAATGACCTGTGC 969
 Db 317 -----IleYsAspHisValIleuAspIyrAsnGlnAlaIleSerArgProHis 331
 QY 970 CCTGTGTCACAGAGGTGTGACAGAAAGCGAGACGCCCTTCATGTCACCTGTGGCGAC 1029
 Db 332 ArgHisMetAspIyrTyrlGlyProSerAspLeuCysPro-----ProIyValProIlys 349
 QY 1030 TCTGCTCACTTCAGAGCAGACAGACAAAGGCCGTGCATCCCGCAGCGTGTCCAGCGTG 1089
 Db 350 SerAlaSerSerAsnCyValPheProGlySerThrHisAspGluSerSerProSerSer 369
 QY 1090 CCCAGGCTGAGCGCGCTGTTAAGACACAGCCCGCTGTGACAGCGCTCAAGTGGCGTG 1149
 Db 370 ProSerValSerGluGlySerAlaAlaSerGluPro----- 381
 QY 1150 CACCTGTCCGACAGACAGCGTGGAGACAGCAATTAAGCTCAAGCGTCTCTGTGAGAT 1209
 Db 381 ----- 381
 QY 1210 ATCAATCAGTTTCATATTTCAGCCGACATGCGAGCATCTTACATGAGCTTCTCTCATCA 1269
 Db 382 -----GluThrSerSerSerAlaIleSer-----SerSerSer 392
 QY 1270 GAAGATGCTTGGAAATCTACAAACTTTCACACTCTGTGATGGACCAACAAGCTATGC 1329
 Db 393 ThrAlaSerAlaProProSerMetProSerThrSerGlnGlnGlyThrSer----- 409
 QY 1330 CAGTTCCTCCCTGTTCAGAACTATCGAGACACATCTCCGAAACCACT----- 1377
 Db 410 -----SerGlyThrValAsnValaIleuGlyIysArgAsnMetThrMetAspLeuGlyLeu 427
 QY 1378 CCGATTAAGAGGAGAGCC--AGCATCCCAAGAGACTGACAGCCGCG-----AGG 1425
 Db 428 ProHisArgProIySalAlaLeuGlyLeuProSerArgIleGlyThrSerValAlaGluLeu 447
 QY 1426 CTTTCAGACAGC--CAGAGCAAGCAATGTGCAATTCGGTCAAGAACAGACAGAGTGGCAC 1482
 Db 448 ProSerProSerThrGluLeuSerArgIyrGluSerPheAsnGlyProGluAlaIleAlaPr 467
 QY 1483 GCCCAGAGGTCCTTTATCTCACTGCATCGAAGTGGAGCGTGGAGACAAATTAACAC 1542
 Db 467 oSerThr--ProIleu-----AsnPheThrAs 476
 QY 1543 ACCAGCTCTCTTTGGCCTTTCACACAGCAGACAGCACTCAGAAAGTGCtGGCGCTG 1602
 Db 476 nProCySphAsnSerProIleIleProValAlaIleSerSerSerArgGluValIle----- 494
 QY 1603 GGCTTAAAGGGCTGGCACTCGGATATCTGGCCCCCAAGACTTTCACCTTCCCTGAC 1662


```

Db      371  ThrmEtaPleuGlueProHleAtrProLysAlaLeuGlueProSerArgIleGly 390
Oy      1417  ACCGCC-----AGCCTTCAGACAGC---CAGAGCAAGCATTTGATTCGGTCGA 1464
Db      391  ThrsrValAlaGlueProSerProSerThrGlueSerArgLeuSerHe-AsnG1 410
Oy      1465  ACCAGCAGCATGGCCAGCCGCCAGAGAGTCCCTTTATCTCCACTGCATGAAAGTGAGC 1524
Db      410  yProGluaAlaAlaAtrProSerThr---ProIleLeu----- 421
Oy      1525  GTGAGAGCAATTACCAACAGCATTCCTTTTCGGCTTTCACCCAGCAGACAGCACTC 1594
Db      422  -----AsnPherThrAsnProCyAspHeAsnSerProIleProValAlaSerSerSe 439
Oy      1585  ACCAAGTCCTGCTGGCGGCTTAAGGCTGACCTGCATATCTTGGCCCCCAGACC 1644
Db      439  rArgGluaValIle----- 443
Oy      1645  TCTACCCCTTCCTGACCAAGCATGCTGATTTTGGCAGAGATCCTGACATTTACTCT 1704
Db      444  -LeuThrLeuProThrProAlaAla-----SerSerSerSerSerThrSe 458
Oy      1705  GCCTCAGCCA-----TCTACGAGGAGCAGTGCAGT 1734
Db      458  rSerGluproSerPheAspPheSerSerPheGlueSerSerSerSerSerIleValVa 478
Oy      1735  TACTCTGCTTACAGCTGACGACGACGCTGCCACTTGGCAGACCAAGCTTATTCGCGC 1794
Db      478  lGluaSerProPhePheAlaSerThrGluaAlaProAlaGlySerSerSerIleSerThrPr 498
Oy      1795  AGCGCGAGCAAGCATGACAGCATGCTGCTGGCGGAGCTGGCATGAAGAGAGCCCC 1854
Db      498  oSerGly---SerIleSerThrProAlaSerAlaSerSerSerAlaSerArgCyAr 517
Oy      1855  TTTGAAA 1861
Db      517  gMeclys 519

```

RESULT 15

```

O7SZF3  PRELIMINARY; PRT; 367 AA.
AC 07SZF3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RX MEDLINE=2388257; PubMed=12477932.
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., McEwan N.A., Peters G.J., Abramson R.D., Mallaby S.J.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Falcetti J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Green E.D., Dickson M.C.,
RA Krzyzanski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Maitra M.A.;

```

```

RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC052477; AAH52477.1; -.
KW Hypothetical protein.
SQ
SEQUENCE 367 AA; 40538 MW; F89979A415DAD6AC CRC64;

Alignment Scores:
Pred. No.: 8 11e-29
Score: 487.50 Length: 367
Percent Similarity: 53.82% Matches: 114
Best Local Similarity: 33.53% Mismatches: 69
Query Match: 13.49% Indels: 124
DB: 13 Gaps: 33

US-10-029-345A-108_copy_538_2532 (1-1995) x Q7SZF3 (1-367)
Oy      43  TTGGTGCTCTGCTGGAAGTGAAGCAAAAGTCTGCTAATTGATAGCCGCACTT 102
Db      15  LeuysArgLeuMetLysAspArgLysCysLeuLeuAspCyArgSerPhe 34
Oy      103  GTGGATACAAATCATCCCAATTTTGGCAAGCATTAATATCACTGCTCCAGCTTATG 162
Db      35  LeuAlaPheSerIleAlaLysLeuArgLysValAlaValAlaLysValAlaValAla 54
Oy      163  AACGCAAG-----TTGCACAGCAGCAAGATGTTAATTTACAGCTCATCAG 210
Db      55  ArgArgArgAlaLysGlySerValSerLeuAspGlnIleLeuSerGlyAsp----- 71
Oy      211  CATTCAGCAAAATVADAGTGGATGACATTGATTCAGTCAAGAGGTTGATTCAGTCA 270
Db      72  AspGluAlaArgSerArgLeuLysSerGlyLeuThrValSerLeuValIleLeuVal 91
Oy      271  AGCTCCCAAGATGTTGCTCTCTCTTCAGAGCTGTTTCACTGATCTTGGGTAAA 330
Db      92  ArgSerSerAspThrAsnThrMetLysAspSerThrIleThrLeuValHisAlaAla 111
Oy      331  CTG---GAGAAGAGTTCAACTC---CTTCACCTGCTTGCAGGTGGTTGCTGATTC 384
Db      112  LeuCyAspArgPheThrPheSerThrGluValLysLeuLysGlyLysArgPhe 131
Oy      385  TCTGCTGTTTCTGCTGCTCTGTG-----GAGGAAAA 417
Db      132  SerThrGlnTyrProAspTyrCysLeuLysThrArgThrLeuSerValSerSerGln 151
Oy      418  TCACTCAGTCCCTACCTGATCTTTCAGACCTTTCCTGTTGCCAATTCAGGCA 477
Db      152  SerSerMetGluSerSerCysLeuSerCysAlaThrProGlnHisAspGlnIlePro 171
Oy      478  ACCGAAATTTCTCCCAATCTTTATCTTGGCTGCGCAGCAGATGCTCTCAAGAGAGCTG 537
Db      172  ValGluIleLeuProPheLeuPheLeuGlySerAlaLeuHisAlaSerLysLysAspPhe 191
Oy      538  ATACAGAGAAATGGATTTGTTGTTAAATGCCAGCTATACCTTCGCAAGCCTGAC 597
Db      192  LeuAspArgMetCylLysSerAlaLeuLeuValSerSerAsnCyProAsn---His 210
Oy      598  TTTATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657
Db      211  PheGluGlyAspTyrGlnTyrLysCysIleProValGluAspAsnHisLysGluAspIle 230
Oy      658  TTGCGGTGTTGGAACAAATCAGTATTCATTTGAAGAAAGCAAGCTCAATGATGT 717
Db      231  SerSerThrPheIleGluAlaIleGluPheIleAspSerValLysAspSerAsnGlyArg 250
Oy      718  GTTTAGTGCATGTTTATGCTGAGATCTCCGCTCCGCGCAACATCGGTATCCGCTCATC 777
Db      251  ValLeuValHisCysGluAlaGlyLysSerArgSerAlaThrIleCysLeuAlaTyrLeu 270

```



```

Oy      778  TGAAGAGACATGCAATGCTTTAATAGAGCTTACATATTTGTAAAGAAAAGAACT  837
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::
Db      271  MetLysIbLysArgValArgLeuIuGlnAlaPheGluPheValLysGlnArgSer  290
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::
Oy      838  ACTATATATCTCCAACTTCATATTTCTGGGCCAACTCTGACATATAGAGAAGATTAAG  897
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::
Db      291  IleIleSerProAsnPheSerPheMetGlyGlnLeuGlnPheGluSerGlnValLeu  310
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::
Oy      898  AACCAAGATGAGACATCAAGGCCCAAGAAGCAAACTCAAGCTGCTGCACCTGGAGAAGCCA  957
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::
Db      311  --AlaIlnSerCysSerValGlnAlaAlaSer----- 320
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::
Oy      958  AATGAACCTGTCCTCGCTGCTCAAGAGGTGAGCAAGAAAGCAGCCCTCAGTCCA  1017
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::
Db      321  -----ProSerAlaSerLeuGlyProLysSerSerSerThrProIlnSerPro  336
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::

```

Search completed: June 21, 2004, 12:40:18
Job time : 138.154 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:40:31 ; Search time 94.9018 Seconds
(without alignments)
11869.448 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 3615
Sequence: 1 atggcccatgagatgatgtg9.....tggaatcatctgagctctcc 1995

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1163542 seqs, 28213646 residues

Total number of hits satisfying chosen parameters: 2327084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/USPTO.epool.p/US10029345/runat.21062004.122818.4331/app.query.fasta.1.3278
-DB=Published Applications AA -QFMT=fastaan -SUFFIX=rabb -MINMATCH=0.1
-LOOPCR=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human4.cdt -LIST=45 -DOCALIGN=200 -THR SCORE=pcr -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MATEL=200000000 -USER=US10029345 @CGN_1_1_21 @runat.21062004.122818.4331
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOF=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppa/PCR_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppa/PCRUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	3406	94.2	665	9	US-09-816-494-2	Sequence 2, Appl1
2	3406	94.2	665	9	US-09-964-277-2	Sequence 2, Appl1
3	3406	94.2	665	12	US-10-072-012-680	Sequence 680, App
4	3406	94.2	665	12	US-10-168-506-14	Sequence 14, Appl
5	3406	94.2	665	15	US-10-343-357-7	Sequence 7, Appl1
6	3406	94.2	665	12	US-10-377-072-26	Sequence 26, Appl1
7	3406	94.2	665	16	US-10-257-026-2	Sequence 2, Appl1
8	3406	94.2	665	16	US-10-648-599-240	Sequence 240, App
9	3406	94.2	665	16	US-10-648-599-247	Sequence 247, App
10	3406	94.2	665	12	US-10-072-012-679	Sequence 679, App
11	3406	94.2	665	12	US-10-072-012-703	Sequence 703, App
12	3406	94.2	665	12	US-10-425-114-54204	Sequence 54204, A
13	3399	94.0	665	12	US-10-072-012-681	Sequence 681, App
14	3399	94.0	665	15	US-10-094-749-2312	Sequence 2312, Ap
15	3379.5	93.5	662	12	US-10-072-012-258	Sequence 258, App
16	3358.5	92.9	660	12	US-10-072-012-256	Sequence 256, App
17	3322	91.9	672	12	US-10-296-115-1259	Sequence 1259, Ap
18	3079.5	85.2	660	12	US-10-072-012-682	Sequence 682, App
19	2954.5	81.7	677	12	US-10-072-012-683	Sequence 683, App
20	2594	71.8	517	9	US-09-964-277-21	Sequence 21, Appl
21	1326	36.7	625	12	US-10-072-012-689	Sequence 689, App
22	1302	36.0	663	12	US-10-072-012-700	Sequence 700, App
23	1297	35.9	253	15	US-10-108-260A-4872	Sequence 4872, Ap
24	1103.5	30.5	616	12	US-10-072-012-266	Sequence 266, App
25	917	25.4	501	12	US-10-072-012-702	Sequence 702, App
26	807	22.3	155	9	US-09-964-277-7	Sequence 7, Appl1
27	735.5	20.3	461	12	US-10-072-012-701	Sequence 701, App
28	654.5	18.1	169	14	US-10-346-356-15	Sequence 15, Appl
29	654.5	18.1	170	9	US-09-775-925-26	Sequence 26, Appl
30	654.5	18.1	170	9	US-09-847-519A-11	Sequence 11, Appl
31	654.5	18.1	170	12	US-10-655-073-16	Sequence 16, Appl
32	654.5	18.1	170	14	US-10-314-058-14	Sequence 14, Appl
33	654.5	18.1	170	14	US-10-405-808-16	Sequence 16, Appl
34	619	17.1	155	9	US-09-964-277-6	Sequence 6, Appl1
35	619	17.1	155	9	US-09-955-732-6	Sequence 6, Appl1
36	469	13.0	444	12	US-09-964-899-47	Sequence 47, Appl
37	469	13.0	442	12	US-10-058-270A-130	Sequence 130, App
38	469	13.0	482	14	US-10-346-356-2	Sequence 2, Appl1
39	469	13.0	482	16	US-10-648-599-164	Sequence 164, App
40	466.5	12.9	381	14	US-10-184-832-2	Sequence 2, Appl1
41	455	12.6	394	9	US-09-736-457-805	Sequence 805, App
42	455	12.6	394	9	US-09-902-941-805	Sequence 805, App
43	455	12.6	394	9	US-09-849-626-805	Sequence 805, App
44	455	12.6	394	12	US-10-283-017-805	Sequence 805, App
45	455	12.6	394	14	US-10-017-754-805	Sequence 805, App

ALIGNMENTS

RESULT 1
US-09-816-494-2
Sequence 2, Application US/09816494
Patent No. US20020034807A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OF INVENTION: 38592 AND 21117, NOVEL DUAL SPECIFICITY
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-09-816-494-2

Alignment Scores:
Pred. No.: 5.84e-234 Length: 665
Score: 3406.00 Matches: 663

Percent Similarity: 99.85%
 Best Local Similarity: 99.70%
 Query Match: 94.22%

Conservative: 1
 Mismatches: 1
 Indels: 0
 Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x US-09-816-494-2 (1-665)

QY 1 ATGGCCCATGAGATGATTTGGAACCAATTTGTTACTGAGAGAGTTGGCTGCTGCGGAA 60
 DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
 QY 61 AGTGAACGMAAAAGTGTCTTAATTGATAGCCGCGCAATTTGGGAATACAAATCATCC 120
 DB 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnThrsmThSer 40
 QY 121 CACATTTGGAGCGCATTTATATCACTGCTCCAGCTTATGAGCGAAGTTGGCAACAG 180
 DB 41 HisIleLeuGlnIleAlaIleAsnIleAsnGlySerLysLeuMetLysArgArgLeuGln 60
 QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTAAGGAAACATAAGTTGATGAT 240
 DB 61 AspLysValLeuLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspLys 80
 QY 241 TGCAGTCAGAAAGTGTAGTTAGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
 DB 81 CysSerGlnLysValValValValLysAspGlnSerSerGlnAspValAlaSerLeuSer 100
 QY 301 GACTGTTTCTCACTGTACTTCTGGGTAACCTGAGAAAGAGCTTCACTGTTCACTG 360
 DB 101 AspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPheAsnSerValHisLeu 120
 QY 361 CTTCAGAGTGGGTGGTGGAGTCTCGCTGTTTCCCTGCTGCTGTGAGGAAATCC 420
 DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgPheProGlyLeuGlyLysSer 140
 QY 421 ACTGAGTCCCTACCTGCAATTTCTCAGCCTTGTACTGTTGCTGCAACATTTGGGCAAC 480
 DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATTCCTCCCAATCTTTATCTTGGCTGCAAGGATGCTCTCAACAGAGAGCTGATA 540
 DB 161 ArgIleLeuProAsnLeuLysLeuGlyCysGlnArgAspValLeuAsnLysGlnLeuMet 180
 QY 541 CAGCAAAATGGATTTGTTATGTGTTAAATGCCAGCTATACCTGTCGCAAGCTGACTT 600
 DB 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
 QY 601 ATCCCGAGTCTCATTTCCGCGGTGCTGATGACAGCTTTGTGAGAAATTTTG 660
 DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
 QY 661 CCGTGTGAGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCANTGATGTGT 720
 DB 221 ProThrPheAspLysSerValAspPheIleGlnLysAlaLysAsnGlyCysVal 240
 QY 721 CTAGTGCACTGTTTGTAGCTGAGATCTCCGCTCCGCAACATCGCTATGCGCTACATG 780
 DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerLalaThrIleAlaIleAlaTyrIleMet 260
 QY 781 AAGAGAGTGAAGATCTTTAGATGAAGCTTACGATTGTGAAAGAAAGAAAGAACTACT 840
 DB 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlnLysArgProThr 280
 QY 841 ATATCTCCAAACTCAATTTCTGGGCAACTCTGGAATATGAGAAAGATTAAGAAC 900
 DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysLysIleLysAsn 300
 QY 901 CAGACTGAGGATATGAGGCGCAAGAGCAAACTCAAGCTGTCGACCTGAGAAAGCAAT 960
 DB 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuLysLysProAsn 320
 QY 961 GAACCTGTCCCTGCTGTCTCAGAGGGTGAACAGAAAGGAGAGCGCCCTCAGTCCACCC 1020

DB 321 GluProValProAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerProPro 340
 QY 1021 TGTGGCGACTGTCTTACCTCAAGGCGACAGCAAGAAAGCCCGTGATCCGCGACGGTG 1080
 DB 341 CysAlaAspSerSerLalaThrSerGlnAlaIleGlnArgGlnArgProValHisProAlaSerVal 360
 QY 1081 CCCAGCGTCCCGAGCGTACAGCCGCTGCTGTATGAGAGCAAGCCCGCTGTACAGCGCTC 1140
 DB 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerAsnLysLeuLysArgSerPhe 380
 QY 1141 AGTGGCGTCAACCTGTCCGAGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
 DB 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 400
 QY 1201 TCTGTGATATCAATCAGTTTCAATTTACGCGAGATGCGAGCATCTTACATGGCTTC 1260
 DB 401 SerLeuAspLysSerValSerTyrSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
 QY 1261 TCTCATCAGAGAGATGCTTGGAAATCAATCAAACTTCCACTACTGTGATGGAGCAAC 1320
 DB 421 SerSerSerGlnLysPheAlaLeuGlnLysTyrLysProSerThrThrLeuAspGlyThrsm 440
 QY 1321 AAGCTATGCAAGTCTCCCTGTTGAGAACTATGCGAGAGACATCCCGAAACAGTCTCT 1380
 DB 441 LysLeuGlyGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
 QY 1381 GATTAAGAGAGAGCGACGATCCCGCAAGAGCTGACAGACCGCGACGCTTCAGACAGCAG 1440
 DB 461 AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 QY 1441 AGCAAGCATTTGATTCGCTCAGAACACAGACAGAGTGGACACGCGCGAGAGTCCCTTTA 1500
 DB 481 SerLysArgLeuHisSerValArgThrSerSerSerGlnThrAlaGlnArgSerLeuLeu 500
 QY 1501 TCTCCACTGATCGAAGTGGAGGCTGGAGAGCAATTAACACACAGCTTCTTTTCGCG 1560
 DB 501 SerProLeuHisAlaGlySerGlySerValGlnAspAsnLysThrSerPheLeuPheGly 520
 QY 1561 CTTTCCACCAAGCAGACAGACATCAAGAGCTGCTGCGCGCTTAAAGGCTGGCAC 1620
 DB 521 LeuSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuLysGlyLysThrHis 540
 QY 1621 TCGATATCTTGGGCCCCCGACAGCTTACCCCTTCCCTGACAGACGCTGTATTTTGGC 1680
 DB 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrPheAla 560
 QY 1681 ACAAGTCTTCACTTCTACTGCTCAGCCATCATGAGGAGGAGGAGGCTGACTTCT 1740
 DB 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
 QY 1741 GCTTACAGCTGCAAGCGCTGCCACTTTCGGGAGACCAAGTCTATTTGTGTGCGACAGCG 1800
 DB 581 AlaTyrSerCysSerGlnLeuProThrCysGlnAspGlnValTyrSerValArgArg 600
 QY 1801 CAGAAAGCAAGTCAAGAGCTGACTGCGCGCGAGCTGAGCAATGAAGAGAGCCCTTGA 1860
 DB 601 GlnLysProSerSerArgAlaAspSerArgAspSerThrPheGlnGlnLysSerProPheGln 620
 QY 1861 AAGCAGTTTAAACGCAAGAGCTGCCAATTTGGAAGAGACATCATGTGAGAGAAC 1920
 DB 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
 QY 1921 AGGTCAAGGAAAGAGCTGGGGAAGTGGGAGTCAAGTCAAGCTTTTGGGAGACATGAA 1980
 DB 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
 QY 1981 ATCATTTAGAGTCTCC 1995
 DB 661 IleIleGlnValSer 665

RESULT 2
 US-09-964-277-2
 ; Sequence 2, Application US/09964277

[illegible]

QY	CTAATGCACTGTTAGCGGGATATCCCGCTCCGCAACATCGGTATCGCTACATCATG	780
Db	LeuValHisCyb5eLysuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlIleMet	260
QY	AAGAAGATGACATGTCCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT	840
Db	LysAlaThrArgAspMetSerLeuAspGluAlaTyrlArgPheValIlySerGluLysArgProThr	280
QY	ATATCTCCAAATTCATATTTTCTGGGCAACTCCTGCACTATGAGGAAGAATTAAGAC	900
Db	IleSerProAsnPheAsnPheLeuGlyIleLeuLeuAspTyrGluLysLysIleLysAsn	300
QY	CAGACTGAGCATCAGAGGCCAAAGAGAAATCAAGCTGCTGCACCTGGAGAACCAAT	960
Db	GlnThrGlyAlaSerGlyProlYsSerIlybLeuLysbLeuLysbLeuGluLysPheAsn	320
QY	GAACCTGTCCCTGCTGTCTCAGAGGGGTGACAGAAAGCCAGACGCCCTCAAGTCCACC	1020
Db	GluProValProAlaValaIleSerGluGlyIleGluLysSerGluThrPheLeuSerProbo	340
QY	TGTCGCCGACTGTGCTACTCAGAGGACGACGAGCAAAAGCCCGGCATCCGCGACGCTG	1080
Db	CysAlaAspSerAlaThrSerGluAlaAlaGlyIleThrGluProValHisPheIleSerVal	360
QY	CCGACGCTGCCCAACCTGCACGCCGCTGCTTTAGAGACAGCCGCTGTGTACAGCCGCTC	1140
Db	ProSerValProSerValGlnProSerIleuLeuGluAspSerPheLeuValGluAlaLeu	380
QY	AGTGGGCTGCACCTGTCCGACAGACAGGCTGGAAAGACAGCAATAGCTCAAGCCTTCTTC	1200
Db	SerIlyLeuHisIleuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe	400
QY	TCTCTGGATATCAATCAGATTCAATTATTCAGCCGACATGGCAGCATCCTTACATGCTTC	1260
Db	SerIleuAspIleuLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe	420
QY	TCCTCATCAGAGATGCTTGGAAATACACAACTTCCACTACTCTTGGATGGAGCAAC	1320
Db	SerSerSerGluAspAlaAlaLeuGluTyrTyrIlyAspProSerThrThrLeuAspGlyThrAsn	440
QY	AAGGTATGCCAGTCTCCCGGTTCAGAGAACTATCCGAGACGACTCCCGGAACACAGTCT	1380
Db	LysLeuLysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro	460
QY	GATAGAAGAGAGCCAGCATCCCAAGAACCTGCAGACCCGCGAGCCTTCACAGACCCAG	1440
Db	AspArgGlnGluAlaSerIlePheProLysbLeuGlnThrAlaArgProSerAspSerGln	480
QY	AGCAAGGATTTGCATTGGGTACAAACCAAGCAGAGTGGACCCGCCAGAGGTCCCTTTTA	1500
Db	SerIlyAspArgLeuHisIleSerValArgThrSerSerSerGlyThrAlaGlnArgSerIleuLeu	500
QY	TCTCCATCGATCGAAGTGGAGGTGGAGAGCAATTAACAACACAGCTTCTCTTTCCGC	1560
Db	SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly	520
QY	CTTTCCACCAAGCAGACGACCTTACGAAAGTCTGCTGCGCTTGAGGCTTGAC	1620
Db	LeuSerThrSerGlnGlnHisIleLeuThrLysSerSerAlaGlyLeuGlyLeuLysGlyTyrHis	540
QY	TGGATATCTTTGGCCCCCAGACCTTACACCTTACCCCTGACACAGACGTGGTATTTTGGC	1680
Db	SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrTyrPheAla	560
QY	ACAGAGTCTTCACATTTACTCTGCTGCCTCAGCCCATCTACGAGGACGAGTCCAGTTACTCT	1740
Db	ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyIlySerAlaSerTyrSer	580
QY	GCCATACGCTGACACGCTGCCCATCTTGGGAGAACCAAGCTATTTCTGTGCGGAGGCG	1800
Db	AlaIlySerCyb5SerGlnLeuProThrCyb5LysAspGlnValTyrSerValArgArgArg	600

QY	901	CAAGCTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTTGAGAAAGCAAAAT	960
Db	301	GlnthrtgylalaserGlyProlysSerIylsleuylsleuethlsleuGlnLylsPAsn	320
QY	961	GAACCTTGCCCTGGCTGCTCAGAGGGGTGACAGAAAAGCSGAGAGCCCTCAGTCCAAACC	1020
Db	321	GluProValProAlaValSerGlnGlyGlyGlnIylsSerGlnThrProIleuSerProPro	340
QY	1021	TGTGGCCACTGTGCTCACTCAGAGGACGACAGCAAAAGCCCGTGCATCCCGCAAGGTG	1080
Db	341	CysAlaIlePserSerAlaThrSerGlnAlaIaIaGlyGlnArgProValIleProAlaSerVal	360
QY	1081	CCCAAGCGTGGCCAGCGTGCAGCCGCTGTTAAGAGACACCCGCTGGTTCAGAGCGCTG	1140
Db	361	ProSerValProSerValGlnProSerIleuGlnIleuPserProIleuValGlnAlaIleu	380
QY	1141	AGTGGCGTGCACCTGTCCGACAGACAGGCTGGAGAGACAGCAATTAAGCTCAACGTTCTCTC	1200
Db	381	SerGlyIleuIlelsleuSerAlaAspArgIleuGlnIleuPserIleuLylsArgSerPhe	400
QY	1201	TCTCTGGATATCAAAATCAAGTTTCATATTCAAGCCAGCATGGAGCATCTTTCATAGGCTTC	1260
Db	401	SerIleuAspIleuylsSerValSerIylsSerAlaSerMetAlaIaSerIleuIleGlyPhe	420
QY	1261	TCTTCATCAGAAAGATGCTTTGAATATCAACAAACCTTCCATCTAGTGGAGACCAAC	1320
Db	421	SerIleSerSerGlnIleuAspAlaIleuGlnLylTrIylsProSerThrThrIleuAspGlyThrAsn	440
QY	1321	AAGCTATGCCAGTTCTCCCTGTTCCAGAACTATGCGAGACAGACTCCGAAACCAAGTCTT	1380
Db	441	LylsleuCyGlnIlePheSerProValGlnIleuSerGlnGlnThrProGlnIleuThrSerPro	460
QY	1381	GATTAAGAGAGAAAGCAGCATCCCAAGAAAGCTGACACCGCAGGCTTCAGACAGCCAG	1440
Db	461	AspIylsGlnGlnIleuAlaSerIleProLylsIylsIleuGlnIleuAlaArgProSerAspSerGln	480
QY	1441	AGCAAGCGATTTGCAATTCGGTCAGAAACCAAGACAGAGTGACACCGCCAGAGGCTCCCTTTTA	1500
Db	481	SerIylsArgIleuIlelsleuSerValArgThrSerSerSerGlyThrIaIaIleArgSerIleuIleu	500
QY	1501	TCTCCACCTGCATCCAAAGTGGGAGCGTGGAGACAATTACACACCAAGTCTCTTTTGCGC	1560
Db	501	SerProIleuIlelsArgSerGlySerValGlnIleuAspIleuIleuIleuIleuSerPheIleuPheIyl	520
QY	1561	CTTTCACACCAACCAAGCAGACCACTCAACGAAGCTGTGCTGCCTTGAGCCTTAAAGCGTGGAC	1620
Db	521	IleuSerThrSerGlnGlnIleuIlelsleuThrIylsSerAlaGlyIleuGlyIleuLylGlyIleuIle	540
QY	1621	TCGGATATCTTTGGCCCCCCCAAGCCTCTACCCCTTCCGTACACAGACGCTGGATATTTGGC	1680
Db	541	SerAspIleuAlaIleProGlnIleuThrSerThrProSerIleuIleuIleuSerIleuIlePheAla	560
QY	1681	ACAGAGCTCTCAACCTTCACTACCTGCGCTCAGCAACTTACGGAGGACAGGCCAGTTAACTCT	1740
Db	561	ThrIleuSerSerHisPheIylsSerAlaSerAlaIleIleIleIleIleIleIleIleIleIleIleIle	580
QY	1741	GCCCTACAGCTCAGCCAGCTGCCCACTTGCCGAGACCAAGTCTATTTCTGTGCGCAGCGCG	1800
Db	581	AlaIylsSerCySerSerGlnIleuProThrCySerGlyAspGlnIleuIleIylsSerValIleArgArgyl	600
QY	1801	CAGAAAGCCAAATGACAGAGCTGACCTCGCGGGGAGAGCTGGCATGAAAGAAAGCCCTTTGAA	1860
Db	601	GlnIylsProSerAspArgAlaAspSerArgIylsSerIleuIlePheIleuIleuIleuIleuIleu	620
QY	1861	AAGCAGTTTAAACGACGAAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATCTCAGAGAAC	1920
Db	621	LylsGlnIleuIylsArgIylsSerCySerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu	640
QY	1921	AGGTACAGGGAAGAGCTGGGAAAGTGGGCACTAGTCTAGCTTTTCGGGACGACATGGAA	1980
Db	641	ArgSerIleArgIleuGlnIleuGlyLylsValGlySerGlnIleuSerIleuSerGlyIleuMetCylu	660

OY		1981	ATCATTTGAGGTCTCC	1995	
Db		661	TlellegluValser	665	
		RESULT 4			
		US-10-168-506-14			
		Sequence 14, Application US/10168506			
		Publication No. US20040053229A1			
		GENERAL INFORMATION:			
		APPLICANT: PLOWMAN, GREGORY D.			
		APPLICANT: MARTINEZ, RICARDO			
		APPLICANT: WHYTE, DAVID			
		APPLICANT: MANNING, GERARD			
		APPLICANT: SUDARSANAM, SUCHA			
		APPLICANT: HILL, RON			
		APPLICANT: FLANAGAN, PETER			
		TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES			
		FILE REFERENCE: 038602/1351			
		CURRENT APPLICATION NUMBER: US/10/168,506			
		CURRENT FILING DATE: 2002-06-21			
		PRIOR APPLICATION NUMBER: PCT/US00/34736			
		PRIOR FILING DATE: 2000-12-21			
		NUMBER OF SEQ ID NOS: 76			
		SOFTWARE: PatentIn Ver. 2.1			
		SEQ ID NO 14			
		LENGTH: 665			
		TYPE: PRF			
		ORGANISM: Homo sapiens			
		US-10-168-506-14			
		Alignment Scores:			
		Pred. No.: 5,84e-234		Length: 665	
		Score: 3406.00		Matches: 663	
		Percent Similarity: 99.85%		Conservative: 1	
		Best Local Similarity: 99.70%		Mismatches: 1	
		Query Match: 94.22%		Indels: 0	
		DB: 12		Gaps: 0	
		US-10-029-345A-108_COPY_538_2532 (1-1995) x US-10-168-506-14 (1-665)			
OY		1	ATGGCCCCATGGAGTGAATTGGAACCTCAATTGTTACTGAGAGGTGGTGCTCGTAGAA	60	
Db		1	Meclahlsiglumetlleglythrctimilevalinrgluayrlevualaaleuleugln	20	
OY		61	AGTGGAAACGAAAAAGTCSTCTAATTGATAGCCGCCCAATTGTGGAATAACAATCATCC	120	
Db		21	Serqlyrhrgluylvalleuleuenileampersearqprophelvalglutrydanthiser	40	
OY		121	CACATTTTGAAGCCATTAAATATGCACCTGCTCCAGCCTTATGAGAGGAAGTTGCACACG	180	
Db		41	Histilleugluialialeamilleamncyssetrlylemetelyargtryglinglin	60	
OY		181	GACAAAGTGTAAATTACACAGCTCATTCAGACATTCAGGAAACATPAGGTTGACATTGAT	240	
Db		61	AapryvaVlaelluethtghluenulleglmhserralalyshlyvalasprileamr	80	
OY		241	TGCAGTCAGAAGTGTAGATTACGATCAAAGCTCCAAAGATGTTGSCCTCTCTCTTGA	300	
Db		81	Cysserglulyvalvalvaltyrhparginsserrginamrvalalasertleuserr	100	
OY		301	GACTGTTTTCTCACTGTACTCTTGSGGTAACGTGAGAAAGCTTCACTCTGTCACTCG	360	
Db		101	AaprcysphelautnrtValleuleucglylysleuglulysserPhabamserValihisleu	120	
OY		361	CTTGCAAGTGGGTTTGTGAGTTCTCTGCTGTTTCCCTGGGCTCTGGAAGAAAATCC	420	
Db		121	LenualeglylphenaleaglrpheserargCywphreogilyLeucybsldubtulysser	140	
OY		421	ACTCTAGACCCTACTCGCATTTCTCAGCCTTGCTTCACTGTGTCCAACTTTGGGCCAAC	480	
Db		141	ThrieuvalprotnrCyaisleserdlnrprocybleupovalalalsanilleglyprotnr	160	
OY		481	CGAAITTTCCCACATCTTTATCTTGGCTGCCAGCAGATGTCCTCAACAAGAGCTGATA	540	

Db 161 ArgIleuProAsnLeuYrleuGlyCysGlnAlaAspValLeuAsnLysGlyTyrPheHis 180
QY 541 CAGCAGAAATGGGATTTGGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 600
Db 181 GlnGlnAsnGlyLleGlyYrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCCGCGTGTGCTGGAATGACAGTTTGTGGAAAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuThrGlyAlProValAsnMetSerPheCysGlnLysIleLeu 220
QY 661 CCGTGTGGACAAATCAGTACATTTCAATTGAGAAAAGCAAAAGCCTCAATGATGTGT 720
Db 221 ProTyrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTATGCACTGTTAGTGGGATCCCGCTCCGCAACCTGATCCGTATCCGTACATCATG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGATGGACATCTTTTATGATGACCTTACAGATTTGTGAAAAGAAAAGCCTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlyAlaTyrArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAAACTTCATTTTCTGGGCCAATCCTGCACTATGAAAGAAATTAAGAAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuAsnAspTyrGlnLysLysIleLysAsn 300
QY 901 CAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAAAGCAAA 960
Db 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlnLysProAsn 320
QY 961 GAACCTGCTCCGCTGTCTCAGAGGGTGGACAAAGAAAGCAGAGCCCTCAGTCCAGCC 1020
Db 321 GluProAlaProAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerProPro 340
QY 1021 TGTGCCACTGTGTACTCAGAGGACAGAGCAAAAGCCGTGATCCGCGCAGCTTG 1080
Db 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGCTGGCCGCGGTGACCCGCTGCTTTAGAGACACAGCCCGTGTGACAGGCGCTC 1140
Db 361 ProSerValProSerValGlnProSerLeuLeuGlnLysSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGCACTGTCCGACAGAGCTGGAAGACAGCAAAATTAACCTCAAGCTTCC 1200
Db 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnLysSerAsnLysLeuLysArgSerPhe 400
QY 1201 TCTGTGATATCAAAATCACTTTCAATTTAGCCAGCATGGACATCCCTTACATGCTTC 1260
Db 401 SerLeuAspIleLysSerValSerLysSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
QY 1261 TCTCATCAGAAAGATGCTTTGGAATACTCAAACTTCCACTCTGGATGGAGCAAC 1320
Db 421 SerSerSerGlnAspAlaLeuGlnLysTyrLysProSerThrThrLeuAspGlyThrAsn 440
QY 1321 AAGCTATGCGACTTCTCCCTGTTCAAGAACTATTCAGAGCAGACTCCGAAACCACTCT 1380
Db 441 LysLeuCysGlnPheSerProValGlnLysLeuSerGlnGlnThrProGlnThrSerPro 460
QY 1381 GATAAGAGAGAAAGCAGAGATCCCAAGAAAGTGCAGACCCGACGCTTCAAGACGCA 1440
Db 461 AspLysGlnGlnLysLeuIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
QY 1441 AGCAAGCATGTCATTGGTGCAGAACAGCAGCAGTGGCAGCCGCGCAGAGTCCCTTTTA 1500
Db 481 SerLysArgLeuHisSerValArgThrSerSerSerLysThrAlaGlnAspSerLeuLeu 500
QY 1501 TCTCCACTGCATCGAAGTGGAGGTGGAGAGCAATTAACAACAAGCTTCTTTTGCGC 1560
Db 501 SerProLeuHisArgSerGlySerValGlnAspAsnTyrHisThrSerPheLeuPheGly 520
QY 1561 CTTTTCACACAGCAGCAGCACTCAGCAAGTCTGTGCTGGGCTTAAAGGCTGGCCAC 1620

Db 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheHis 540
QY 1621 TGGGATATTTGGCCCCCAAGACTTACCCCTTCCCTGACACAGCAGCTGATTTGGC 1680
Db 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
QY 1681 ACAGAGTCCCTACACTTACTCTGCGCTCAGCAGCATCTACAGGAGCACTGGCACTTACT 1740
Db 561 ThrLysSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
QY 1741 GCTTACACTGACGACGACCTGCCCACTTGGCGAGACCAAGTCAATTTCTGTGCGCAGCG 1800
Db 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValAlaArgAlaArg 600
QY 1801 CAGAAAGCAAGTGCACAGAGCTGACTCGCGCGGAGAGCTGCGATGAAGAGACCCCTTTGAA 1860
Db 601 GlnLysProSerAspAlaAspSerArgArgSerTyrPheGlnLysSerProPheGln 620
QY 1861 AAGCATTTAAACGACAGACCTGCCAAATGGAATTTGAGAGAGCATCATGTACAGAAC 1920
Db 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlnSerIleMetSerGluAsn 640
QY 1921 AGTCCAGGGAGAGCTGGGGGAAAGTGGGACAGTCACTTACCTTTTGGGCGACATGGAA 1980
Db 641 ArgSerArgGlnGlnLeuGlnLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATTGAGTCTCC 1995
Db 661 IleIleGlnValSer 665

RESULT 5
US-10-343-357-7
Sequence 7, Application US/10343357
Publication No. US20040058341A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
APPLICANT: BILLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
APPLICANT: YAO, Monique G.; BURROUD, Neil
APPLICANT: WANG, Yumel E.; STEWART, Elizabeth A.
APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
APPLICANT: LEE, Ernestine A.; HAFALIA, April J.A.
APPLICANT: LU, Dying Aina M.; TRIBOULET, Catherine M.
APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
APPLICANT: YUE, Henry; WARREN, Bridget A.
APPLICANT: NGUYEN, Daniel B.; CHAMLA, Nandinder K.
APPLICANT: KEARNEY, Liam
TITLE OF INVENTION: PROTEIN PHOSPHATASES
FILE REFERENCE: PI-0173 PCT
CURRENT APPLICATION NUMBER: US/10/343,357
PRIOR FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,679
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/223,272
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/224,309
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 60/226,728
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/229,254
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: US 60/231,366
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CD1

Db 661 llellegluValSer 665

```
RESULT 6
US-10-377-072-26
; Sequence 26, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Miyoung
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/10/377,072
; CURRENT FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 665
; TYPE: PRP
; ORGANISM: Homo Sapiens
US-10-377-072-26
```

```
Alignment Scores:
Pred. No.: 5,84e-234
Score: 3406.00
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 94.22%
Matches: 665
Conservative: 1
Mismatch: 1
Indels: 0
Gaps: 0
```

US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-377-072-26 (1-665)

```
QY 1 ATGGCCGATGAGGATTTGGAACCTCAATTTGTAACGAGAGTTGGTGGCTGCTGAGAA
Db 1 MetalaHsGluMeIlleGlyThrglnlIeValThrglnrgeuValAlaleuLeuGln 20
QY 61 AGTGAACGAGAAAAGTGTCTTAATTGATAGCCGCGCAATTTGTGGAATACATATCATCC
Db 21 SerGlyThrglnrgeuValleuLeuIleAspSerArgProPheValGlnrYasnThrsr 40
QY 121 CACATTTGGAAGCATTTATATCACTGCTCCAAAGCTTATAGAGGAAGGTTGCAAG
Db 41 HistleuGlnlAlaleuIleAenIleAenCySerIyLeuMeIlyArgrgrleuGlnIn 60
```

```
QY 181 GACAAAGTGTAAATACAGAGCTCATCCAGATTCAAGCAAAATAGTTGACATTGAT
Db 61 AspLySValleuIlleThrglnleuIlleGlnHsSerAlaYshIlySValAspIleAsp
QY 241 TGCAGTCAGAAAGTTGTGTTTACATCAAAAGCTCCCAAGATGTTGCTCTCTTCA
Db 81 CySerGlnIySValValValAlaIlyrAspIleSerIeGlnAspValAlaSerLeuSer
QY 301 GACTGTTTCTCACTGTATCTTGGGTAACTGGAGAAAGACTTCAACTGTTCACG
Db 101 AspCySpheLeuThrValleuLeuGlyLySleuGlnIySerIeHsSerValHsleu
QY 361 CTTCAGAGTGGATTGGAGGTTCTCCGTTGTTTCCCTGCGCTGTGAGAAATCC
Db 121 leuAlaGlyGlyPheAlaGlnPheSerArgCySphProGlyLeuCySglnGlySer
QY 421 ACTGTATCCCTTACCTTATCTTCAAGCTTGTACGCTTGTACTGTTCACAACTGGCCAC
Db 141 ThrleuValProThrCySleSerGlnProCySleuProValAlaAsnIleGlyProThr
QY 481 CGAATCTCCCAATCTTATCTTGGCTGCGCAGGATGTCCTTAACAGAGCTGATA
Db 161 ArgIleleuProAsnIeulrleuGlyCySglnArgAspValleuAsnIyGlnleuMet
QY 541 CAGCAAAATGGAGTTGTTATGTTAAATGCACGCTATACCTGTCCAAAGCTGACTT
Db 181 GlnGlnAsnIyIleGlyrValleuAsnIaSerAenThrCySProLySProAspPhe
QY 601 ATCCCGAGTCATATTCCTGCGTGTGCTGGAATGACACTTTTGTGAAAAATTTG
Db 201 IleProGlnSerHsIlePheleuArgValProValAsnAspSerPheCySglnIySleu
QY 661 CCGTGTGTGCAAAATCGATGATTTCAATTGAGAAAGCAAAAGCTCAATGATGTT
Db 221 ProThrleuAspLySserValAspPheIleGlnIySAlaIySAlaSerAnGlyCySVal
QY 721 CTAGTCACTGTTAGCTGGAGATCCCGCGCCACCATCGCTATCGCTTACATCATG
Db 241 leuValHsIyCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAGAAAGAAAGCTT
Db 261 LysArgMetAspMetSerleuAspGlnAlaTyArgPheValIySleuIySArgProThr
QY 841 ATATCTCCAAACTCATTTTCTGGGCAACTCTGACTATGAGAAAGATTAGAAC
Db 281 IleSerProAsnIleAsnIleuGlyGlnleuLeuAspTyrglnIySleIySleYasn
QY 901 CAGACTGAGATCAGAGGCAAAAGCAAACTCAAGCTGTGACCTGAGAGCAAT
Db 301 GlnThrglnAlaSerGlyProLySserIySleuIySleuHsIleuGlnIySProAsn
QY 961 GAACCTGCTCCCTGCTGTCAAGAGGTTGACGAAAGAGAGGAGGAGGAGGAGGAGG
Db 321 GluProValProAlaIleSerIyGlnIyGlnIySserGlnIyTrProLeuSerProPro
QY 1021 TGTGCCGACTGTCTTACTCAAGGCGACAGCAAAAGCCGTGACATCCGCGCAGGCTG
Db 341 CyAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHsProAlaSerVal
QY 1081 CCCAGGTCCTCCAGGCTGAGCGCGTGTATAGAGACAGCCGCTGTGACAGGCGCTC
Db 361 ProSerValProSerValGlnProSerIeueuGlnAspSerProleuValGlnAlaLeu
QY 1141 AGTGGGCTGACCTTCTCGCAGACAGAGGCTGAGACAGCAATAGCTCAAGGTTCTTC
Db 381 SerGlyLeuHsIleuSerAlaAspArgleuGlnAspSerIySleuIySArgSerPhe
QY 1201 TCTCTGATATCAATCATGTTATATTCAGGAGATGGCAGACATCTTACATGCTTC
Db 401 SerleuAspIleIySserValSerIySerAlaSerMetAlaIaIaSerleuHsIyGlyPhe
QY 1261 TCCATCATAGAAAGTCTTTGAAATCAACAACTTCACTACTGTGATGGACCAAC 1320
```

```

Db      421 SerSerSerGlnuPhaAlaLeuGluTyrTyrLeuProSerThrThrLeuMetGlyThrAsn 440
QY      1321 AACGATGCGAGTTCTCCCTGTTACAGAACTATCGAGCAGCACTCCGAAACCACTGCT 1380
Db      441 LysLeuGlnuPhaSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
QY      1381 GATTAAGAGAAAGCAAGCACTCCCAAGAACTGCAAGCCGCGCTTACAGACGCG 1440
Db      461 AspLysGlnGlnuAlaSerLysLeuGlnThrAlaArgProSerAspSerGln 480
QY      1441 AGCAAGGATTTGATTCGATCGATCAGAACGAGCAGCAGGACCGCCAGAGGTCCTTTTA 1500
Db      481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
QY      1501 TCTCCATGTCATCGAAGTGGAGCGTGGAGAGCAATTACCAACCAAGCTTCTTTTGCGC 1560
Db      501 SerProLeuHisArgSerGlySerValGlnAspAsnTyrHisThrSerPheLeuPheGly 520
QY      1561 CTTTCCACCGCCAGCAGCACTTCACGAAATCTGCTGGCCTGGGCTTAAAGGCTGGCAG 1620
Db      521 LeuSerThrSerGlnGlnuHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPhe 540
QY      1621 TCGGATATCTGGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGATATTTGCC 1680
Db      541 SerPheLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
QY      1681 ACAGAGTCTCTACACTTCTGCTGCTGACCACTTACGAGGAGGAGTGCAGTTACTCT 1740
Db      561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAlaSerTyrSer 580
QY      1741 GCCTACAGCTGCGACCGCTGCCACTTGGGAGAGCAACTCTATCTGTCGCGACGCGG 1800
Db      581 AlaTyrSerCysSerGlnuLeuProThrCysGlyAspGlnValTyrSerValAlaArgArg 600
QY      1801 CAGAAGCGCAAGTACAGAGCTGACTCGCGGCGAGCTGCATGAAGAGAGCCCTTTGAA 1860
Db      601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrPheGlnuSerProPheGln 620
QY      1861 AAGCAGTTTAAACGACAGACCTGCGCAATGAAATTGGAGAGAGCATGTCAGAGAAC 1920
Db      621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyLysSerIleMetSerGlnuAsn 640
QY      1921 AGGTACAGGGAAGAGCTGGGGAAGTGGGAGATCAGTCTGCTTTTGGGCGACAGTGA 1980
Db      641 ArgSerArgGlnuGlnuLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY      1981 ATCATGAGGTCTCC 1995
Db      661 IleIleGlnuValSer 665

RESULT 7
US-10-029-345a-108_copy_538_2532.rapb
; Sequence 2, Application US/10257026
; Publication No. US20040086859A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10KDM5
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-257-026-2

Alignment Scores:
Pred. No.: 5,84e-234 Length: 665
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1

```

```

Best Local Similarity: 99.70% Mismatches: 1
Query Match: 94.22% Indels: 0
DB: 16 Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-257-026-2 (1-665)

QY      1 ATGGCCCATGAGATGATGGAATCAATGATTTACTGAGAGGTTGGCTGCTGTCGAA 60
Db      1 MetAlaHisGlnuMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
QY      61 AGTGAACGGAAGAAAGCTGCTAATGATTAAGCGGACCATTTGTGAATACATACATCC 120
Db      21 SerGlyThrGlnuLysValLeuLeuIleAspSerArgProPheValGlnuThrSer 40
QY      121 CACATTTTGAAGCCATTATATCACTGCTCCAACTTATGAAAGGAAAGTTGCAACG 180
Db      41 HisIleLeuGlnuAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGlnGln 60
QY      181 GACAAAGTGTAAATTAACAGAGCTCATCCAGATTGCGGAAACATTAAGTTGACATTGAT 240
Db      61 AspLysValLeuIleThrGlnuLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY      241 TGCAGTCAGAAAGTTGATTAACATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCA 300
Db      81 CysSerGlnuLysValValValTyrAspGlnSerSerGlnuAspValAlaSerLeuSer 100
QY      301 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAAGAGCTTCAACTCTGTTCACTG 360
Db      101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY      361 CTTGCAAGTGGGTTTCTGAGTTCTCTGTTGTTTCCCTGCGCTCTGTGAGGAAATCC 420
Db      121 LeuAlaGlyGlyPheAlaGlnuPheSerArgCysPheProGlyLeuGlyLysSer 140
QY      421 ACTGATGCTCTACCTGATTTCTGACCTTCTGCTTCACTGTTGCAATGGGCAACC 480
Db      141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAlaAsnIleGlyProThr 160
QY      481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGACGAGATGTCCTCAACAAAGAGCTGATA 540
Db      161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGlnuMet 180
QY      541 CAGCAGAAATGGAGTTGTTATGTTAAATGCCAGGATATACCTGTCAAGCTGACTTT 600
Db      181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY      601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG 660
Db      201 IleProGlnuSerHisPheLeuArgValProValAsnAspSerPheCysGlnuLysIleLeu 220
QY      661 CCGTGTGGAACAATCAAGTATGATTTTCAATTGAGAAAGCAAAAGCTTCAATGATGTGT 720
Db      221 ProThrLeuAspLysSerValAspPheIleGlnuLysAlaLysAlaSerAsnGlyCysVal 240
QY      721 CTAGTGCATGTTTACTGGAATCTCCGCTCGGACCAACATGCTATGCTATGCTATCATG 780
Db      241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY      781 AAGAGATGAGACATGCTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGACTTACT 840
Db      261 LysArgMetAspPheSerLeuAspGlnuLysArgPheValLysGlnuLysArgProThr 280
QY      841 ATATCTCCAAACTTCAATTTTCTGGGCGCAATCTTGACATTAAGAAAGAAAGTAAAGAAC 900
Db      281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysIleLysAsn 300
QY      901 CAGACTGAGCAATCAGGCGCAAGAGCAAACTCAAGCTGTGACCTGGAAGAGCAAAAT 960
Db      301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLysLeuLysLeuLysProAsn 320
QY      961 GAACCTGTCTGCTGCTCTCAGAGGGTGAACAGAAAGCGAGAGCGCCCTCACTGACACC 1020
Db      321 GluProValProAlaValSerGlnuGlyGlnuLysSerGlnuThrProLeuSerProPro 340

```

QY 1021 TGTGCGGACTGTGCTTCACTCAGAGCAGCAGACAAAGCCCGCTGATCCCGCAGCGTG 1080
 Db 341 CysAlaAspSerAlaThrSerGlnAlaIleGlnAlaGlyProValHisProAlaSerVal 360
 QY 1081 CCCAGGTCGCCAGCGTGTGAGCCGCTGCTGTTAGAGACGCCCGCTGTGACAGCGGCTC 1140
 Db 361 ProSerValProSerValGlnProSerIleuIleuIleuIleuIleuIleuIleuIleu 380
 QY 1141 AGTGGGCTGCACCTGTCCGAGACAGGTGTGAGACAGCAATTAAGCTCAAGCTTCCTTC 1200
 Db 381 SerGlnIleuHisIleuSerAlaAspArgIleuGlnAspSerAsnIleuIleuIleuIleu 400
 QY 1201 TCTCGATATCAAAATCAATTCATATTATTCAGCCAGCATGTGACAGCATTCCTTAAGCTTC 1260
 Db 401 SerIleuAspIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 420
 QY 1261 TCTCATCAGAAAGATGTGCTTGGAAATCAAAACCTTCACATCTGTGATGGACCAAC 1320
 Db 421 SerSerSerGlnAspAlaIleuGlnIleuIleuIleuIleuIleuIleuIleuIleuIleu 440
 QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAGAACTATGCGAGACAGACCTCCGAAACCAAGTCTT 1380
 Db 441 LysIleuGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 460
 QY 1381 GATTAAGGAGGAAAGCCAGCATCCCAAGAAAGCTGCGACCGCCAGGCTTCAGACAGCAG 1440
 Db 461 AspIleuGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 480
 QY 1441 AGCAAGCGATTCGATTCGGCTCAAGAACAGCAGCAGTGTGACCGCCGAGAGTCCCTTTA 1500
 Db 481 SerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1500
 QY 1501 TCTCACTGTCATGAGAGTGTGAGCGTGTGAGAGCAATTAACACACAGCTTCCTTTGCGC 1560
 Db 501 SerProIleuHisIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 520
 QY 1561 CTTTCACACGCGCAGCAGCACTCAAGCTGTGCTGCTGCGCTTAAAGGCTGTGAGCAG 1620
 Db 521 LeuSerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 540
 QY 1621 TCGATATCTTGGGCCCCCAGACCTCAACCTTCCTTCCTCAACGAGCTGTATTTGGC 1680
 Db 541 SerAspIleuAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1680
 QY 1681 ACAGAGCTTCACACCTTCTACTCTGCTCAAGCATCTAAGAGCAGCAGTGTGAGTTCCT 1740
 Db 561 ThrGlnSerSerHisPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 580
 QY 1741 GCTCAAGCTGTGAGCAGCAGTGTGAGCAGTGTGAGCAGCAAGCTATTTCTGTGCGAGCG 1800
 Db 581 AlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1800
 QY 1801 CAGAAAGCGAGTGTGAGCAGTGTGAGCAGCAGTGTGAGCAGTGTGAGCAGCAGCCTTTGAA 1860
 Db 601 GlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1860
 QY 1861 AAGCAATTTAAACGAGAGCTGTGAGCAGCAGTGTGAGCAGCAGCAGCAGCAGCAGCAGCAG 1920
 Db 621 LysGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1920
 QY 1921 AGGTCAAGGAGAGAGCTGTGAGCAGCAGTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1980
 Db 641 ArgSerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1980
 QY 1981 ATCATTTAGGCTCTCC 1995
 Db 661 IleIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 660

RESULT 8
 US-10-648-593-240
 ; Sequence 240, Application US/10648593
 ; Publication No. US20040106132A1

; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
 ; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
 ; FILE REFERENCE: D0273 NP
 ; CURRENT APPLICATION NUMBER: US/10/648,593
 ; PRIOR APPLICATION NUMBER: 60/406,385
 ; NUMBER OF SEQ ID NOS: 557
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO: 240
 ; LENGTH: 665
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-648-593-240
 Alignment Scores:
 Pred. No.: 5,846-234
 Score: 3406.00
 Percent Similarity: 99.85%
 Best Local Similarity: 99.70%
 Query Match: 94.22%
 Ds: 16
 Gaps: 0
 US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-648-593-240 (1-665)
 QY 1 ATGGCCCATGATGATGTTGAACTCAATTTGTTACTGAGAGTGTGCTGTGCGAGAA 60
 Db 1 MetAlaHisGlnMetIleGlyThrGlnIleValIleThrGlnIleuValAlaIleuIleuGln 20
 QY 61 AGTGAACGGAAGAAAGTCTGCTAATGATGATGCGGCGCATTTGTGGAATCAATATCATCC 120
 Db 21 SerGlyThrGlnIleuValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 40
 QY 121 CACATTTTGAAGGCAATTAATATATCACTGCTCCAGCTTATGAAGGAGTGTGCAACAG 180
 Db 41 HisIleuGlnIleuAlaIleuAsnIleuAsnIleuAsnIleuIleuIleuIleuIleuIleu 60
 QY 181 GACAAAGTGTAAATTAACAGAGCTCAATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
 Db 61 AspIleuValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 240
 QY 241 TGCAAGTGAAGAGTGTGATTTGCATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
 Db 81 CysSerGlnIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 300
 QY 301 GACTGTTTCTCACTGATCTTCTGGGTAAACTGAGAGAGCTTCAACTCTGTTCACTG 360
 Db 101 AspCysPheIleuThrValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 360
 QY 361 CTTCAGAGTGTGAGTGTGCTGATTTCTGTTGTTTCCCTGCGCTGTGAGAGAAATCC 420
 Db 121 LeuAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 420
 QY 421 ACTCTATGCTCACTGATTTCTCAGCTTGTGCTTACCTGTGCAACATTTGGCGCAACC 480
 Db 141 ThrIleuValProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 480
 QY 481 CGAATTTCTCCCAATCTTATCTTGTGCTGCTGCGCAGAGAGATCTCTCAACAGAGAGTGATA 540
 Db 161 ArgIleuIleuProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 540
 QY 541 CAGCAGAAATGGAGTGTGATTTGTTAAATGCGCAGCTATACCTGTGCCAAAGCTGACTTT 600
 Db 181 GlnGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 600
 QY 601 ATCCCGAGTCTCAATTTCCGCGTGTGCTGTGATGACAGCTTTTGTGAGAAAAATTTTG 660
 Db 201 IleProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 660
 QY 661 CCGTGGTTGGAACAATCAATGATTTTCATTGAGAAAGCAAAAGCTCCAAATGATGTGTT 720


```
QY 361 CTTCAGAGTGGGTTGCTGAGTCTCTGCTGTTTCCCTGAGCTCTGTGAGAGAAATCC 420
Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlySer 140
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGGCTTGCTTACCTGTGCAACATTTGGGCCAAC 480
Db 141 ThrLeuValProThrCysEilSerGlnProCysLeuProValAlaAlaSerIleGlyProThr 160
QY 481 CGAATTTCTTCCCAATCTTATCTTGGCTGCCAGGAGATGCTCTCAAGAGAGAGCTGATA 540
Db 161 ArgIleLeuProAlaSerLeuArgValCysGlnArgValLeuSerIleGlyLeuMet 180
QY 541 CAGCAGATGGGATTTGTTATGTGTAAATGCCAGCTATACCTGTCCAAAGCTTCACTTT 600
Db 181 GlnGlnAenGlyIleGlyTyrValIleuAlaSerAlaThrCysProIleProAspPhe 200
QY 601 ATCCCGAGTCTCAATTTCTGCGTGCCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAlaAspSerPheCysGlnIleLeu 220
QY 661 CCGGTGTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAGCTTCCAAATGATGTT 720
Db 221 ProIlePheAspIleSerValAspPheIleGlnIleValIleAlaSerAlaSerVal 240
QY 721 CTAGTCACTGTTTACCTGAGATCTCCGCTCCGCAACATGCTTATGCTTACATCATG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTGTGAGAAAGAAAGAAAGCTTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnIleTyrArgPheValIleGlnIleValAspProThr 280
QY 841 ATATCTCCAAATTTCTGAGGCAACTCTGCACTATGAGAAAGATTTTAAAGAC 900
Db 281 IleSerProAlaPheAlaPheLeuGlyGlnLeuAspIleTyrGlnIleValSerVal 300
QY 901 CAGATGAGAGATCAGGCGCCAAAGAGCAAACTCAAGCTGTGCACTGAGAGAGCGCAAT 960
Db 301 GlnThrGlyAlaSerGlyProIleSerIleLeuIleLeuIleLeuIleLeuIleLeuIleLeu 320
QY 961 GAACCTGCTCCCTGCTCTCAGAGGGTGGACAGAAAGAGAGAGCGCCCTCAGTCCACC 1020
Db 321 GlnProValProAlaValSerGlnIleGlyGlnIleValSerGlnIleProIleSerProPro 340
QY 1021 TGTCGCACTGTGCTACCTCAGAGGCAAGAGCAAAAGCGCTGATCCGCGCAAGGCTG 1080
Db 341 CysAlaAspSerAlaThrSerGlnIleAlaGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGCTGTGCCAGCGTGCAGCGCTGTGAGAGAGACCGCGCTGTGACAGGCGCTC 1140
Db 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGCACTGCTCCGCAAGAGGCTGAGAGAGCAAGCAATAGCTCAGAGCGTTC 1200
Db 381 SerGlyLeuHisIleuSerAlaAspArgLeuGlnAspSerAlaSerLeuHisIleGlyPhe 400
QY 1201 TCTCTGATATCAAAATCAGTTTCATATTCAGCGAGAGATGAGAGATCTTCAACATG 1260
Db 401 SerLeuAspIleLeuSerValSerIleSerIleSerIleSerIleSerIleSerIleSerIle 420
QY 1261 TCTCATCAGAGAGAGCTTTGGATATCTCAAACTTCCATCTCTGATGGAGAGCAAC 1320
Db 421 SerSerSerGlnAspAlaLeuGlnIleTyrIleValSerProThrThrLeuAspGlyThrIle 440
QY 1321 AAGCTATGCCAGTTCTCCCTGTGTGAGAGACTATCGGAGAGAGCTCCGAAACAGTCTT 1380
Db 441 LysLeuCysGlnPheSerProValGlnGlnIleuSerGlnIleProGlnIleSerPro 460
QY 1381 GATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 461 AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480

QY 1441 AGCAGAGATTCATTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 481 SerIleArgLeuHisSerValIleArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
QY 1501 TCTCCATCTGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 501 SerProLeuHisArgSerGlySerValGlnAspSerHisThrSerPheLeuPheGly 520
QY 1561 CTTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 521 LeuSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuGlyLeuLysGlyTrpHis 540
QY 1621 TCGGATATCTTGGCCCCCAGAGAGCTTACCCCTTCCCTGACAGAGAGAGAGAGAGAG 1680
Db 541 SerAspIleLeuAlaProGlnIleThrSerThrProSerLeuThrSerSerTrpIleAla 560
QY 1681 ACAGAGTCTTCACACTTCTACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
QY 1741 GCTTACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValIleArgArg 600
QY 1801 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db 601 GlnLysProSerAspArgAlaAspSerArgAspSerTrpHisGlnGlnIleSerProPheGln 620
QY 1861 AAGCAGTTTAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db 641 ArgSerArgGlnGlnIleuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATTTGAGGCTCC 1995
Db 661 IleIleGlnValSer 665

RESULT 10
US-10-072-012-679
/ Sequence 679, Application US/10072012
/ Publication No. US2004003493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tchernov, Vellizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Ratturajan, Meera
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangolli, Beba
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Coleman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grose, William M.
/ APPLICANT: Alsbrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Kieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072, 012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
```

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remining prior application data removed - See File Wrapper or PAM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 679
LENGTH: 690
TYPE: PR1
ORGANISM: Homo sapiens
US-10-072-012-679

Alignment Scores:
Pred. No.: 5,89e-234 Length: 690
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Beet Local Similarity: 99.70% Mismatches: 1
Query Match: 94.22% Indels: 0
DB: 12 Gaps: 0
US-10-029-345a-108_COPY_538_2532 (1-1995) x US-10-072-012-679 (1-690)
QY 1 ATGGCCCATGATGATTTGGAATCTGTAAGAGGTTGGCTGCTGCGAA 60
DB 26 MetAlaHisGluMetCileGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 45
QY 61 AGTGAACGGAAGAAAGCTGCTAATGATAGCGGCGATTGGTGAATCAATACATCC 120
DB 46 SerGlyThrIuLysValLeuLeuIleAspSerArgProPheValGluIlyrAsnThrSer 65
QY 121 CACATTTTGAAGCATTATATCACTGCTCAAGCTTATGAAGCGAAGTTGCAACAG 180
DB 66 HistIleuLunIalIeAenIleAsnCySerIlySeuMetIySArgIleuGlnGln 85
QY 181 GACAAAGTTTAATTACAGAGCTTCACGATTCAGCGAAACATNAGTTGACATTGAT 240
DB 86 AspIyValIleuIleThrGluLeuIleGlnHisSerAlaIySValAspIleAsp 105
QY 241 TGAAGTGAAGGTTGATGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB 106 CysSerIlnySValValIyTyAspGlnSerSerIlnAspValAlaSerIleuSer 125
QY 301 GACTGTTTCTCACTGATCTTGGAGTAACAGGAGAGCTTCACTGTTCACTG 360
DB 126 AspCySHeuHeuThrValLeuLeuGlyIyLeuGluIySerPheAsnSerValHisLeu 145
QY 361 CTTCGACAGTGGTTTGGTAGTTCTCTGTTGTTCCCTGAGCTCTGTAAGAAATCC 420
DB 146 LeuAlaIyGlyPheAlaGluPheSerArgCySHeProGlyLeuCySglIyIySer 165
QY 421 ACTTACTCCCTACCTGATCTTCAGCCTTCGATCTGTTGCAACATGGGCAAC 480
DB 166 ThrIeuValProThrCySHeSerGlnProCySHeuProValAlaAsnIleGlyProThr 185
QY 481 CGAATTTCTCCCATCTTTATCTTGCTGCAGCGAGATGTCCTCAACAAGAGCTGATA 540
DB 186 ArgIleuProAsnLeuIyIleuGlyCySglInArgAspValIleuAsnIySglIleuMet 205

QY 541 CAGCAGATGGATGTTATGTTAAATGACAGCTATACCTGCAAGCCTGACTTT 600
DB 206 GlnGlnAsnGlyIleGlyTyValLeuAsnAlaSerAsnThrCySProIySProAspPhe 225
QY 601 ATCCCGAGTCTCATTTCTGCGCTGTGCTGTGAATGACAGCTTTTGTGAAATTTTG 660
DB 226 IleProGlnSerHisPheLeuArgValProValAsnAspSerHecySglIySileu 245
QY 661 CCGTGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTTCAATGATGTT 720
DB 246 ProThrLeuAspIySerValAspPheIleGlyIyAlaIySAlaSerAsnGlyCySVal 265
QY 721 CTAGTGACCTGTTTACTGGGAATCTCCGCTCCGACCATGCTATGCTATCATCATG 780
DB 266 LeuValHisCySHeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIyIleMet 285
QY 781 AAGAGATGAGACATGCTTAGATGAAGCTTACAGATTGTGTGAAGAAAGACCTACT 840
DB 286 IySArgMetAspMetSerIleuAspGluAlaIyTyArgPheValIySglIySArgProThr 305
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGACTATGAGAAAGATTAAAGAC 900
DB 306 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyIyIySlySArg 325
QY 901 CAGACTGAGCATCAGAGGCGCAAGAGCAAACTCAAGCTGTCGACCTGGAGAGCCAAAT 960
DB 326 GlnThrGlyAlaSerIyProIySeryIySleuIyLeuLeuIyIleuGluIySProAsn 345
QY 961 GAACCTGCTCTGCTCTCAGAGGCTGACAGAAAGAGAGAGCCCTCAGTCCACCC 1020
DB 346 GluProValProAlaIySerGluGlyGlnIySeryIyIyProIySeryProPro 365
QY 1021 TGTGCGACTTGTCTACTCAGAGGCGACAGACAAAGCCGCTGATCCGCGAGCTG 1080
DB 366 CySAlaAspSerAlaThrSerGluAlaIyGlnArgProValHisProAlaSerVal 385
QY 1081 CCCAGGTCGCCAGCGCTGAGCGCTGCTTATGAGAGACAGCCGCTGAGAGCGGCTC 1140
DB 386 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProIySValGlnAlaLeu 405
QY 1141 AGTGGCTGCACCTGTCCGACAGACAGCTGAGAGACAGCAATNAGCTCAAGCTTCTTC 1200
DB 406 SerGlyLeuHisIleuSerAlaAspArgLeuGluAspSerAsnIySleuIySArgSerPhe 425
QY 1201 TCTCTGATATCAATCACTTCAATTCAGCCAGCATGCGACATCTTACATGCTTC 1260
DB 426 SerLeuAspIleIySerValSerTySerAlaSerMetAlaIaSerLeuHisGlyPhe 445
QY 1261 TCTCATCAGAAAGTCTTGGAAATCTAACAACTTCCACTTCTGATGGAGCAAC 1320
DB 446 SerSerSerGluAspAlaLeuGluIyTyTyIySProSerThrThrIleuAspGlyThrAsn 465
QY 1321 AAGCTATGCAAGTCTCCCTGTTTCAAGAACTATGCGAGACAGCTCCGAAACCAATGCT 1380
DB 466 IySleuCySglInPheSerProValGlnGluIleuSerGlnGlnThrProGluThrSerPro 485
QY 1381 GATNAGAGAGAACCGACATCCCAAGAAAGCTGACAGCCGACGCTTCAAGACCCAG 1440
DB 486 AspIyGlyGluIyAlaSerIleProIySlyIyLeuGlnThrAlaArgProSerAspSerGln 505
QY 1441 AGCAACGATTCGATTCGTCGAGAACAGAGAGGAGGACCGCCAGAGGTCCTTTTA 1500
DB 506 SerIySArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArgSerLeuLeu 525
QY 1501 TCTCACTCATCGAAGTGGAGCGTGAAGCAATATCACACACAGCTTCTCTTTCGAG 1560
DB 526 SerProIySHisArgSerGlySerValGluAspAsnTyHisIleHisSerPheLeuPheGly 545
QY 1561 CTTCACACAGCAGCAGACCTCAGCAAGTCTGCTGCGCTTGAAGGCTGAGCAC 1620
DB 546 LeuSerThrSerGlnGlnHisIleuThrIySeryAlaGlyIyLeuGlyIyLeuIySglIyTrpHis 565


```
Qy 1621 TCGATATCTTGGCCCCGAGACTTACCCCTTCCCTGACGACAGCTGGTATTTGGC 1680
Db 566 SerAspIleuAblProdiInhrSerThrProSerLeuThrSerSerTrpYrPheAa 585
Qy 1681 ACGAGTCTCTACACTTCTACTCTGCTTCCCTGACGACATTAAGGAGGACATGCTACTCT 1740
Db 586 ThrGluSerSerHisPheThrSerSerAlaSerAlaIleThrGluSerSerAlaSerYrSer 605
Qy 1741 GCTTACAGCTGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 606 AlaYrSerCySerSerGluLeuProThrCySerGluAspIleValYrSerValArgArgArg 625
Qy 1801 CAGAGCCAGTACAGAGCTGACTGCGGCGGAGCTGCGATGAAGAGAGCCCTTTGAA 1860
Db 626 GlnYrSerProSerAspArgAlaAspSerArgArgSerTrpHisGluLeuSerProPheGlu 645
Qy 1861 AAGCATTTAAACCCAGAGCTGCAATGGAATTTGGAGAGCATCATGTCAGAGAAC 1920
Db 646 LysGlnPheYrAspArgSerCySerGlnMetGluPheGluSerIleMetSerGluAsn 665
Qy 1921 AGGTACAGGAGAGAGCTGCGGAGAAAGTGGGACAGTCACTGCTTTTGGGAGAGATGAA 1980
Db 666 ArgSerArgGlnGluLeuGluYrValGluYrSerGlnSerSerPheSerGlySerMetGlu 685
Qy 1981 ATCATTTAGGCTCTCC 1995
Db 686 IleIleGluValSer 690
```

RESULT 11

US-10-072-012-703

Sequence 703, Application US/10072012

Publication No. US2004003493A1

GENERAL INFORMATION:

```
APPLICANT: Tcherev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernusen, Bryan
APPLICANT: Patcurajan, Meera
APPLICANT: Shinkens, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Esha
APPLICANT: Padiganu, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
```

```

; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 703
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-703
```

Alignment Scores:

Pred. No.:	Score:	Length:
5,89e-234	3406.00	690
Percent Similarity:	99.85%	Matches: 663
Best Local Similarity:	99.70%	Conservative: 1
Query Match:	94.22%	Mismatches: 1
DB:	12	Indels: 0
		Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-072-012-703 (1-690)

```
Qy 1 ATGGCCCATGATGATGATGGAATCAAAATTTGTTACTGAGAGTTGGTGGCTGCTGGA 60
Db 26 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 45
Qy 61 AGTGAACCGGAAAAGTGGCTGTAATGATAGCCGCGCATTTGGTGAATCAATACATTC 120
Db 46 SerGlyThrGluValValLeuLeuIleAspSerArgProPheValGlnYrAsnThrSer 65
Qy 121 CACATTTGGAAGCATTAATATCACTGCTCCCAAGTTTGAAGCGAAGTTGCAACAG 180
Db 66 HisIleLeuGluValAlaIleAsnIleAsnGlySerIleLeuMetYrAspArgLeuGlnGln 85
Qy 181 GACAAAGTGTATTAATTAACAGCTGATCAGACATTCAGCGAAACATTAAGTTGATCAT 240
Db 86 AspYrValLeuIleThrGluLeuIleGlnHisSerAlaYrHisYrValAspIleAsp 105
Qy 241 TGCAGTCAGAAAGTGTGATTTAGATCAAAAGCTCCCAAGATGTTGCTTCTCTTCA 300
Db 106 CysSerGlnYrValValIleYrAspGlnSerSerGlnAspValAlaSerLeuSerSer 125
Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAAACTGAGAGAAGCTTCACTGCTTCACTG 360
Db 126 AspCyPheLeuThrValLeuLeuGluYrLeuGluYrSerPheAsnSerValHisIleu 145
Qy 361 CTTCAGAGTGGTTTCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
Db 146 LeuAlaGlyYrPheAlaGluPheSerArgCysPheProGlyLeuCySerGluYrSer 165
Qy 421 ACTCTAGTCCCTAACCCTGACATTTCTGAGCTTCTTACCTGTTGGCAATTTGGGCAAC 480
Db 166 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 185
Qy 481 CGAATTTCTCCCAATCTTATCTTGGCGGCGAGGAGATGCTTCAACAAAGAGCTGATA 540
Db 186 ArgIleLeuProAsnLeuYrLeuGlyCysGlnIleAspValLeuAsnYrGluLeuMet 205
Qy 541 CAGGAGAAATGGAGTTGGTTAGTGTAAATGCCAGCTATCCTGCTCAAAAGCTGACTTT 600
Db 206 GlnGlnAsnGlyIleGlyYrValLeuAsnAlaSerAsnThrCysProYrAspPhe 225
Qy 601 ATCCCGAGTCTCATTTCTGCGTGTGCGCGTGAATGAAGAGCTTTTGTGAGAAATTTTG 660
Db 226 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluYrIleLeu 245
Qy 661 CCGTGTGTTGACAAATGAGTGAATTTCAATGAGAAAGCAAAAGCTTCAATGATGTTG 720
Db 246 ProTrpLeuAspYrSerValAspPheIleGluYrAlaYrAlaSerAsnGlyCysVal 265
```


QY 721 CTAGTGACCTTTAGTGGGATCTCCGCTCCGCCACATCGTATCGCTACATCANG 780
DB 266 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 285
QY 781 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTGTGAAAGAAAAGAACTACT 840
DB 286 LysArgMetAspMetSerLeuMetArgIleAlaTyrArgPheValIleGlyIleValProThr 305
QY 841 AATATCTCAAACTTCAATTTTCTGGGCAATCTCTGACATATGAAAGAAAGATTAAAGAAC 900
DB 306 IleSerProAsnPheAsnPheLeuGlyIleLeuLeuAspTyrGlyIleValIleValAsn 325
QY 901 CAGACTGAGACATCAGGCGCCAAAGAAAGAACTGATGCTGACCTGAGGAGAACTCAAT 960
DB 326 GluThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuIleGlyIleValProAsn 345
QY 961 GAACCTGCTCCCTGCTGCTCAGAGAGGTGACAGAAAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 346 GluProValProAlaValSerGlyIleGlyIleLysSerIleThrProLeuSerProPro 365
QY 1021 TGTGCGCAGCTCTGCTACCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 366 CysAlaAspSerAlaThrSerGlyAlaAlaGlyIleArgProValHisProAlaSerVal 385
QY 1081 CCGAGGCTGCGCCAGCGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 386 ProSerAlaProSerValGlnProSerLeuGlnLysPheProLeuValGlnAlaLeu 405
QY 1141 AGTGGGCTGACCTGTCCGAG 1200
DB 406 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysValSerPhe 425
QY 1201 TCTCTGATATCAATAGTTTCAATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 426 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 445
QY 1261 TCTCTATCAAGATGCTTTGGAATATCAAACTTCCATCTCTGATGAGAGAGAGAGAG 1320
DB 446 SerSerSerIleAspAlaLeuGlyIleTyrLysProSerThrThrLeuAspGlyThrAsn 465
QY 1321 AAGCTATGACAGTTCTCCCTGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 466 LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnIleThrProGlnThrSerPro 485
QY 1381 GATTAAG 1440
DB 486 AspLysGlyIleGlnIleAspSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 505
QY 1441 AGCAAGGATTTGATTTGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 506 SerLysArgIleLeuHisSerValArgIleThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 525
QY 1501 TCTCCATGATGAG 1560
DB 526 SerProLeuHisArgSerGlySerValGlnAspAsnTyrHisThrSerPheLeuPheGly 545
QY 1561 CTTTTCACAG 1620
DB 546 LeuSerThrSerGlnGlnIleHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyThrPhe 565
QY 1621 TCGGATATCTGGGCCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB 566 SerAspIleLeuAlaProGlnIleThrSerThrProSerLeuThrSerSerIleTyrPheAla 585
QY 1681 ACAGAGCTCTACACTTACTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
DB 586 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyIleSerAlaAspTyrSer 605
QY 1741 GCTTACAGCTGACAG 1800
DB 606 AlaTyrSerCysSerGlnLeuProThrCysGlyAspAlaIleValTyrSerValAlaArgArg 625

QY 1801 CAGAGCAAGATGACAG 1860
DB 626 GlnLysProSerAspArgAlaAspSerAlaArgSerThrPheGlnGlnLeuSerProPheGln 645
QY 1861 AAGCAGTTTAAACGAG 1920
DB 646 LysGlnPheLysArgIleArgSerCysGlnMetGluPheGlyGlnSerIleMetSerGluAsn 665
QY 1921 AGGTCAAGGAG 1980
DB 666 ArgSerArgGlnGlnLeuGlyIleValGlySerGlnSerSerPheSerGlySerMetGlu 685
QY 1981 ATCATGAGAGTCTCC 1995
DB 686 IleIleGlnValSer 690

RESULT 12
US-10-425-114-54204
; Sequence 54204, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54204
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-028-H6_F11.Dep
; US-10-425-114-54204

Alignment Scores:
Pred. No.: 5, 89e-234 Length: 690
Score: 3406.00 Matches: 663
Percent Similarity: 99.858 Conservative: 1
Best Local Similarity: 99.704 Mismatches: 0
Query Match: 94.224 Indels: 0
DB: 12 Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-425-114-54204 (1-690)
QY 1 ATGGCCCATGAGATGATGGAATGGAATGTTTACTGAGAGGTGGCTGCTGGAA 60
DB 26 MetAlaHisGluMetIleGlyThrGlnIleValIleThrGluArgLeuValAlaLeuLeuGln 45
QY 61 AGTGAACGGAGAAAGAGTGTGCTAATTGATAGCGGGCAATTTGTGGAATTAACAATATCC 120
DB 46 SerGlyThrGlnLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 65
QY 121 CACATTTGGAAGCCCTTATATATCAACTGCTCCAACTTATGAAGGAGAGAGAGAGAGAG 180
DB 66 HisIleLeuGlnAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 85
QY 181 GCAAAAGTATTAATTAACAGAGCTCATCCAGATTCAGGAAACATTAAGGTGACATTGAT 240
DB 86 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaIleValIleValIleAspIleAsp 105
QY 241 TCGAGTCAGAGAGGTGATTAAGATCAAGATCAAGATCCCAAGATGGCTCTGCTTGA 300
DB 106 CysSerGlnLysValValValIleTyrAspGlnSerSerGlnAspAlaAlaSerLeuSerSer 125
QY 301 GACTGTTTCTCAGCTACTTCTGGGTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

```
Db      126 AspCysPheLeuThrValIleuLeuGlyIysLeuGluIysSerPheAsnSerValIhIleu 145
QY      361 CTGGCAGGTGGGTTTGCTGAGATCTCTGTTGTTTCCCTGGCCCTGTGAAAGAAATCC 420
Db      146 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuGlyIysSer 165
QY      421 ACTTATGTCCTTACCTGATTTCTCAGCCTTGCTTACCTGTGCAACATTTGGCCAC 480
Db      166 ThrIleuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 185
QY      481 CGAATTTCTCCCAATCTTTATCTTGCTGCTCCAGCAGAGATGCTCTCAACAGAGAGCTGATA 540
Db      186 ArgIleuProAsnIleuIylLeuGlyCysGlnArgAspValIleuAsnIleGlyLeuMet 205
QY      541 CAGCAGAAATGGGATTTGTTATGTGTTAAATGCGAGCTATACCTGTGCAAAAGCTTACCTT 600
Db      206 GlnIleuAsnGlyIleGlyIylValIleuAsnAlaSerAsnThrCysProIylSerAspPhe 225
QY      601 ATCCCGAGTCTCATTTCCCTGCTGCTGCTGTAATGACAGCTTTTGTGAGAAAATTTTG 660
Db      226 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIylIleu 245
QY      661 CCGGTGTGGACAAATCAGTGAATTTCAATTTGAGAAAGAAAAGCTCCAAATGATGTT 720
Db      246 ProThrLeuAspIylSerValAspPheIleGlyIylValIleuAsnIleGlyCysVal 265
QY      721 CTAGTCACCTGTTTACCTGGATCTCCGCTCCGCTCCGACCATGCTATCGCTACATCATG 780
Db      266 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIylIleMet 285
QY      781 AAGAGATGAGACATGCTTTAGATGAACCTTACGATTTGTGAAAGAAAAGAAACCTACT 840
Db      286 LysAspGluMetAspMetSerLeuAspGluAlaIylArgPheValIylGluIylAspProThr 305
QY      841 ATATCTCCAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTAAGAAC 900
Db      306 IleSerProAsnPheAsnIleuGlyIleuLeuAspIylGlyIylIleAsn 325
QY      901 CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAGCTGTGACCTGTGAGAAAGCAAT 960
Db      326 GlnThrGlyAlaSerGlyProIylSerIylLeuIylLeuIleuIleuIylSerProAsn 345
QY      961 GAACCTGTCCCTGTGCTCAGAGGGGTGACAGAAAAGGAGAACCCCTTCAGTCCACC 1020
Db      346 GluProValProAlaValSerGluIylGlyIleuIylSerGluThrProLeuSerProPro 365
QY      1021 TGTGCGACCTGTCTACCTCAGAGGAGAGCAAAAGGCCCGCTGATCCAGCGCTTC 1080
Db      366 CysAlaAspSerAlaThrSerGluAlaIleGlyIleuArgProValHisProAlaSerVal 385
QY      1081 CCCAGCGTGCAGCGTGCAGCGGTGCTGTAGAGGACAGCCCGCTGTGTAACAGCGCTC 1140
Db      386 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 405
QY      1141 AGTGGGCTCACCTGTCCGACAGAGGCTGGAAGACAGCAATAGCTCAAGCGCTTC 1200
Db      406 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIylSerIylValSerPhe 425
QY      1201 TCTTGATATCAATCAGTTTCAATTAATTCAGCCGACATGGCAGATCTTACATGGCTTC 1260
Db      426 SerLeuAspIleIylSerValSerIylSerIleSerMetAlaIleSerLeuHisGlyPhe 445
QY      1261 TCCGATACGAAGATGCTTTGGAATACTACAAACCTTCACTCTGATGGAGGACCAAC 1320
Db      446 SerSerSerGluAspAlaLeuGluIylIylProSerThrThrIleuAspGlyThrAsn 465
QY      1321 AAGCTATGCAGTTCTCCCTGTTCAGAACTATTCGAGACAGATCTCCGAAACAGTCTCT 1380
Db      466 LysLeuCysGlnPheSerProValGlnIleuSerGluGlnThrProGluThrSerPro 485
QY      1381 GATAAGAGAGAAAGCCAGCATCCCAAGAAAGCTGACAGCCGAGCTTCAGACAGCGAC 1440
Db      486 AspIylGluGluAlaSerIleProIylIylSerGluGlnThrAlaArgProSerAspSerGln 505

QY      1441 AGCAGCGATTTGATTCGATCGATGAGAACAGCAGACGACCGCCAGAGTCCCTTTTA 1500
Db      506 SerIylArgLeuHisSerValAlaArgHisSerSerSerGlyThrAlaGlnArgSerLeu 525
QY      1501 TCTCCACCTGACATCGAAATGGGAGCGGTGAGAGCAATTAACACACAGCTTCTTCCG 1560
Db      526 SerProLeuHisArgSerGlySerValGluAspAsnThrHisThrSerPheLeuPheGly 545
QY      1561 CTTTCACGACGACAGACACCTCAGCAAGATGCTGTGCTGGGCTTAAAGGCTTGCAAC 1620
Db      546 LeuSerThrSerGlnGlnHisLeuThrIylSerAlaGlyLeuIylLeuIylSerGlyThrHis 565
QY      1621 TCGATATCTTGCCCTCCCGCCAGACCTTACCCCTTCCCTGACAGCAGCTGATTTTGC 1680
Db      566 SerAspIleLeuAlaProGlnThrThrSerThrProSerLeuThrSerSerThrIylPheAla 585
QY      1681 ACAGAGTCTTCACACTTCTACTCTGCTGACGATTAAGAGGACAGTGCAGTACTCT 1740
Db      586 ThrGluSerSerHisPheIylSerAlaSerAlaIleIylGlyIylSerAlaSerIylSer 605
QY      1741 GCCTACAGCTGACAGCCAGCTGCCACTTGGCGGAGACCAAGTATATTCTGTGCGAGCGG 1800
Db      606 AlaIylSerCysSerGlnLeuProThrCysGlyAspGlnValIylSerValAlaArgArg 625
QY      1801 CAGAAAGCAAGTACAGACGCTGACTCGCGGAGAGCTGGCATGAAAGAGAGCCCTTTGAA 1860
Db      626 GlnIylProSerAspArgAlaAspSerArgArgSerThrPheIylGluIylSerProPheGlu 645
QY      1861 AAGCAGTTTAAACGAGAAAGCTGCGCAATGGAATTTGGAGAGACATCATGAGAAAC 1920
Db      646 LysGlnPheIylArgArgSerCysGlnMetGluPheGlyIylSerIleMetSerGluAsn 665
QY      1921 AGGTCACGGAAAGACTGGGAAAGTGGGAGTCAAGCTTATTTGGGCGACAGTGA 1980
Db      666 ArgSerArgGluGluLeuGlyIylValGlySerGlnSerSerPheSerGlySerMetGlu 685
QY      1981 ATCATTTGAGGTCTCC 1995
Db      686 IleIleGluValSer 690

RESULT 13
US-10-072-012-681
/ Sequence 681, Application US/10072012
/ Publication No. US2004003493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tchernov, Velizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Paturajan, Meera
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangoli, Beha
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Raselli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Wolenc, Steven D.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grosse, William M.
/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072,012
/ CURRENT FILING DATE: 2002-01-31
```


QY 1621 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGAACAGACAGCTGTATTTTGGC 1680
Db 541 SeraspIleuAlaProGlnThrSerThrProSerIleuThrSerIleuThrPheAla 560
QY 1681 ACAGAGCTCTCAGACTTCTACTCTGCTCAGCCATCTAAGGAGGCGAGTTCCTTACT 1740
Db 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
QY 1741 GCTTACAGCTGAGCGAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCGACGCGG 1800
Db 581 AlaTyrSerTyrSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
QY 1801 CAGAGCGCAAGTACAGAGCTGACTCGCGCGGAGCTGAGCATGAGAGAGCCCTTTGAA 1860
Db 601 GlnIleProSerSerAspArgAlaAspSerArgArgSerTTPHISGlnGlnSerProPheGln 620
QY 1861 AACGAGTTTAAAGCGAGAGCTGCCAAATGAAATTGAGAGAGCATGTCAGAGAAC 1920
Db 621 LysGlnPheLysArgArgSerTyrSerGlnMetGlnPheGlnGlnSerIleMetSerGlnAsn 640
QY 1921 AGGTCAAGGAGAGCTGGGAGAAAGTGGGAGTCACTGCTTCTTGGGAGAGATGGA 1980
Db 641 ArgSerArgGlnGlnGlnGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCAATGAGAGTCTCC 1995
Db 661 IleIleGlnValSer 665

RESULT 14

US-10-094-749-2312
Sequence 2312, Application US/10094749
Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: MAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2312
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-10-094-749-2312

Alignment Scores:

Pred. No.: 1.84e-233 Length: 665
Score: 3399.00 Matches: 662
Percent Similarity: 99.70% Conservative: 1
Best Local Similarity: 99.55% Mismatches: 2

Query Match: 94.02% Indels: 0
DB: 15 Gaps: 0
US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-094-749-2312 (1-665)

QY 1 ATGGCCCATGAGATGATTTGGAATCTCAATTTGTAATGAGAGTTGTGCTGTGGAGAA 60
Db 1 MetAlaHisGlnMetCileGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
QY 61 AGTGAACGGAAAAAGTGTCTGTAATGATGAGCGGCACTTGTGGAATTCATATCATCC 120
Db 21 SerGlyThrGlnLysValIleLeuLeuIleAspSerThrProPheValGlnTyrAsnThrSer 40
QY 121 CACTTTTGAAGCATTAATATATCACTGCTCCAAAGCTTAATGAAGCAAGTTGCAACAG 180
Db 41 HisIleLeuGlnValAlaIleAsnIleAsnTyrSerTyrIleuMetCysArgArgLeuGln 60
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTTGAGCAAACTTAAGTTGACATTGAT 240
Db 61 AspLysValIleuIleThrGlnLeuIleGlnHisSerAlaLysPheValAspIleAsp 80
QY 241 TGCAGTCAGAAAGTGTGATTTGCATCAAGCTCCAAAGATGTTGCTCTCTCTTCA 300
Db 81 CysSerGlnLysValValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTCACTGCTGAGTTCGGGTAACTGAGAGAGCTTCACTCTGTTCACTG 360
Db 101 AspCysPheLeuThrValIleuLeuGlnLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTGAGAGTGGGTTGTTGCTGAGTTCCTGTTGTTTCCCTGGCTCTGTGAGAGAAAAATCC 420
Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlyLysSer 140
QY 421 ACTTATGCTTCACTGCTGATTTCTCAGGCTTGTCTTCACTGTTGCAATTTGGCCCAAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTCCCAACTTATCTTATCTGGCTGCCAGGAGATGTCTCAACAGAGATGATA 540
Db 161 LeuIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuLeuLysGlnLeuMet 180
QY 541 CAGCAAAATGGGATTTGTTATGTATGTTAAATGCCAGCATATCTGTGCAAAAGCTGACTT 600
Db 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnHisSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
QY 661 CCGTGTGGAACAATCAGTAGATTTCAATTGAGAAAGAAAGCTCCCAATGAGATGTTT 720
Db 221 ProThrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTAGTGACAGTTTATGTTGAGATCTCCGCTCGCCAGCATGCTATGCTTACATCATG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGATGACATGCTTATGATGAGCTTACAGATTTGGAAGAAAGAAAGCTTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAAACTCAATTTTCTGGGCAACTCTCTGAGCTATGAGAGAAAGATTAAGAAC 900
Db 281 IleSerProAsnProAsnPheAsnLeuGlyGlnLeuLeuAspTyrGlnLysLysIleLysAsn 300
QY 901 CAGACTGAGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTTGGAGAAAGCAAT 960
Db 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLysLeuGlnLysProAsn 320
QY 961 GAACCTGTCCTGCTGCTCAGAGGCTGAGACAGAAAGAGAGAGCCCTCAGTCCACCC 1020
Db 321 GlnProValProAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerProPro 340

QY 1021 TGTGCCACTCTGCTACTCAGAGGAGGAGCAAAAGCCGCTGATCCCGCAGCGTG 1080
 DB |||||
 QY 341 CysAlaAhpSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 DB |||||
 QY 1081 CCCAGCGTCCCGCAGCGCTGCGCTGTTAGAGGACGCCGCTGTATCAGCGCTC 1140
 DB |||||
 QY 361 ProSerValProSerValGlnProSerLeuGlnAhpSerProLeuValGlnAlaLeu 380
 DB |||||
 QY 1141 AGTGGCGTGCACCTGCTCCGAGACAGGCTGGAAGACAGCAATAAGCTTACGCTTCC 1200
 DB |||||
 QY 381 SerGlyLeuHisLeuSerAlaAlaPargLeuGlnAhpSerHisLeuValArgSerPhe 400
 DB |||||
 QY 1201 TCTCTGGATATCAATTCAGTTTATATTCAGCCAGCATGGCAGCATCTTACATGCGCTTC 1260
 DB |||||
 QY 401 SerLeuAhpLileysSerValSerTyrSerHisSerMetAlaHisSerLeuHisGlyPhe 420
 DB |||||
 QY 1261 TCCTCATCAGAAAGTCTTTGGAATATCTACAAACCTTCCATCTGTGATGGAGCACAC 1320
 DB |||||
 QY 421 SerSerSerGlnAhpAlaLeuGlnTyrTyrTyrProSerThrThrLeuAhpGlyThrAsn 440
 DB |||||
 QY 1321 AACGATGCGCAATTCCTCCCTGTTTCAGGAACATATCGAGAGACATCCCGAAACCATGCTCT 1380
 DB |||||
 QY 441 LysLeuGlySerGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
 DB |||||
 QY 1381 GATTAAGGAGGAGGAGGAGCATCCCGCAAGAGCTGAGACCGCCAGGCTTTCAGACAGCAG 1440
 DB |||||
 QY 461 AspPargGlnGlnAlaSerLileProLysLysLeuGlnThrAlaArgProSerAhpSerGln 480
 DB |||||
 QY 1441 AGCAAGGATGATTCGCTCAGAAACGAGCAGCATGGAGCACCGCCAGAGGTCCTTTTA 1500
 DB |||||
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 DB |||||
 QY 1501 TCTCCCATGATGGAAGTGGAGCGGTGAGAGCAATTCACACACAGCTTCTTTTGCGC 1560
 DB |||||
 QY 501 SerProLeuHisArgSerGlySerValGlnAhpAsnTyrHisThrSerPheLeuPheGly 520
 DB |||||
 QY 1561 CTTTCCACGAGCAGCAGCAGCAGCTCAGAGATGCTGCTGCGGCGCTTAAAGGCTGGCAC 1620
 DB |||||
 QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlnLysGlyTyrPheHis 540
 DB |||||
 QY 1621 TCGGATATCTGGACCCCGCAGACCTTACCCCTTCCCTGACACAGCAGCTGTATTTGCC 1680
 DB |||||
 QY 541 SerAspLileLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPyrPheAla 560
 DB |||||
 QY 1681 ACAGAGTCTCTCAACTTCTACTGCTCAGCCATCTACGAGGAGGAGGAGGAGTACTCT 1740
 DB |||||
 QY 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer 580
 DB |||||
 QY 1741 GCGTACAGTGCAGCGCAGCTGCCCACTTGGGAGAGCAAGTCTATTCTGGCGAGCGG 1800
 DB |||||
 QY 581 AlaTyrSerCysSerGlnLeuProHisCysGlyAhpGlnAlaTyrSerValArgParg 600
 DB |||||
 QY 1801 CAGAAAGCAAGTGAAGAGTGAATCGCGGCGAGGAGTGGCATGAAGAGAGCCCTTTGAA 1860
 DB |||||
 QY 601 GlnLysProSerAhpArgAlaAhpSerArgArgSerTyrHisGlnGlnLysSerProPheGln 620
 DB |||||
 QY 1861 AACGATTTAAAGCAGACACTGCCAAATTTGAATTTGGAAGAGCATCATGTCAAGAAC 1920
 DB |||||
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerLileMetSerGlnAsn 640
 DB |||||
 QY 1921 AGGTCAAGGAGAGGAGGAGGAGGAGGAGGAGTCAAGTCAAGTCTTTTCGCGACATGGA 1980
 DB |||||
 QY 641 ArgSerArgGlnGlnLysValGlyValGlySerGlnSerSerPheSerGlySerMetGln 660
 DB |||||
 QY 1981 ATCATGAGGCTCTCC 1995
 DB |||||
 QY 661 IleIleGlnValSer 665
 DB |||||
 RESULT 15
 US-10-072-012-258
 ; Sequence 258, Application US/10072012
 ; Publication No. US20040033493A1
 ; GENERAL INFORMATION:

; APPLICANT: Tchenev, Velizar
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Zernusen, Bryan
 ; APPLICANT: Patnurajan, Meera
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Li, Li
 ; APPLICANT: Gangoli, Neha
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Anderson, David W.
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Taupier Jr, Raymond J.
 ; APPLICANT: Gusev, Vladimir Y.
 ; APPLICANT: Colman, Steven D.
 ; APPLICANT: Wolenc, Adam R.
 ; APPLICANT: Pena, Carol E. A
 ; APPLICANT: Furtak, Katarzyna
 ; APPLICANT: Grose, William M.
 ; APPLICANT: Alsobrook II, John P.
 ; APPLICANT: Lepley, Denise M.
 ; APPLICANT: Rieger, Daniel K.
 ; APPLICANT: Burgess, Catherine E.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-258
 ; CURRENT APPLICATION NUMBER: US/10/072, 012
 ; FILE FILING DATE: 2002-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,102
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: 60/265,514
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,517
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,412
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,395
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/266,406
 ; PRIOR FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: 60/266,767
 ; PRIOR FILING DATE: 2001-02-05
 ; PRIOR APPLICATION NUMBER: 60/267,057
 ; PRIOR FILING DATE: 2001-02-07
 ; PRIOR APPLICATION NUMBER: 60/266,975
 ; PRIOR FILING DATE: 2001-02-07
 ; PRIOR APPLICATION NUMBER: 60/267,459
 ; PRIOR FILING DATE: 2001-02-08
 ; REMAINING PRIOR APPLICATION data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1391
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 258
 ; LENGTH: 662
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-072-012-258
 Alignment Scores:
 Pred. No.: 4,5e-232 Length: 662
 Score: 3379.50 Matches: 660
 Percent Similarity: 99.40% Conservative: 1
 Best Local Similarity: 99.25% Mismatches: 1
 Query Match: 93.49% Indels: 3
 DB: 12 Gaps: 1
 US-10-029-345A-108_COPY_538_2532 (1-1995) x US-10-072-012-258 (1-662)
 QY 1 ATGGCCCATGAGATGATGGAATGGAATGTTACTGAGAGGTTGGCTGCTGGA 60
 DB |||||
 QY 1 MetAlaHisGlnMetLileGlyThrGlnLileValThrGlnArgLeuValAlaLeuGln 20
 DB |||||
 QY 61 AGTGAACGGAAGAAAGTGTGCTAATGTGATGAGCGGCACTTGTGGAATACATATCCTC 120
 DB |||||
 QY 21 SerGlyThrGlnLysValLeuLeuLileAhpSerArgProPheValGlnTyrAenThrSer 40
 DB |||||

QY	121	CACATTTTGGAAAGCCATTAAATATCACTGCTCCAAAGCTTATGAAGCAAGGTTGCAACAG	187
Db	41	HstIleuGluValAlaIleasnIleasnCySerIysIeuMetIysArgArgLeuGlnGln	60
QY	181	GACAAAGGTGTAATTAACAGAGCTCATCCAGAGATTACAGAAACATAAGTGGTGAATTGAT	240
Db	61	AspIysValIeuIleThrGluIeuIleGlnHisSerAlaIysIleValAspIleAsp	80
QY	241	TGCAGTCGAAGGTTGTAAGTTTACGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA	300
Db	81	CysSerGlnIysValIleValIleAspGlnSerSerGlnIleValAlaIleSerLeuSerSer	100
QY	301	GACGTGTTTCTCACGTGACCTCTGGGTAACATGGAGAAAGGCTTCGACTGTTCACCTG	360
Db	101	AspCysPheLeuThrValIleIleuGlyIysIleuGlnIySerPheAsnSerValHisIleu	120
QY	361	CTTGCAAGTGGGTTTGCTGAGTTCTCGTTGTTTCCCTGGGCTCTGTGTGAAGAAATCC	420
Db	121	LeuAlaGlyIglPheAlaGlnPheSerArgCysPheProGlyIeuCysGlnGlyIySer	140
QY	421	ACTTAGTCCCTACCTGACGATTTCTGACGCTTGCTTACCTGTTCGCAACATTGGGCCAAC	480
Db	141	ThrIleuValProIthrCysIleSerGlnProCysIleuProValAlaAsnIleGlyProIthr	160
QY	481	CGAATTCCTCCAAATCTTTATCTTGGGTGGCCAGACGATGCTCTCAACAAGAGGCTGATA	540
Db	161	ArgIleIleuProAsnIleuTyrlleuGlyCysGlnIleArgAspAlaIleuAsnIySleuMet	180
QY	541	CACAGAGATGGGAAATGGTTATGTTTAAATGCGACGATATCTGTGCCAAAGCCTGACTTT	600
Db	181	GlnGlnAsnGlyIleGlyIyValIleuAsnAlaSerAsnThrCysProIySerProAspPhe	200
QY	601	ATCCCGAGTCTCATTTCCGTGCGTGTGCGCTGATGAATGACAGGCTTTGTGAGAAATTTTG	660
Db	201	IleProGlnSerHisPheIleuArgValProValAsnAspSerPheCysGlnIySleu	220
QY	661	CCGTGGTTGGACAAATCAGTAGATTTTCATTGAGAAAGAAAAGCCTCCAAATGATGTGTT	720
Db	221	ProIyPheAspIySerValAspPheIleGlyIyAlaIyAlaSerAsnGlyCysVal	240
QY	721	CTTAGTCACATGTTTACCTGGGATCTCCGCGTCGCGACCATGCGTATGCGCTTACATATG	780
Db	241	LeuValHisCysIleuAlaGlyIleSerArgSerAlaIleAlaIleAlaIyrlleMet	260
QY	781	AAGAGGATGACATGTCTTTAATGAAGCTTACAGATTTGTGAAAGAAAAGAAAGCCTACT	840
Db	261	LysArgMetAspMetSerIleuAspGlnAlaTyrlArgPheValIyGlnIyIyArgProIthr	280
QY	841	ATATCTCCCAAACCTCAATTTCTGGGGCAAATCCTCGACATATGAGAAAGAATTTAAGAC	900
Db	281	IleSerProAsnIyPheAsnIyGlnIleuIyAspIyGlnIyIyAlaIleIyAsn	300
QY	901	CAGACTGAGCATCAGGGCCAAAGACAAACTCAAGCTGCTGCACCTGGAAGAAGCAAAAT	960
Db	301	GlnThrGlyAlaSerGlyProIySerIySleuIySleuHisIyGlnIyIyProAsn	320
QY	961	GAACCTGTCCTCCGCTGCTCAGAGGGGTGACAGAAAAGCAGACGCGCTTCAGTCCACC	1020
Db	321	GlnProValProAlaValSerIleGlyGlnIyIySerGlnIyIyProIleuSerProPro	340
QY	1021	TGTGCGCACTGTGTAACCTCAGAGGACACAGACAAAGCGCCGCGATCCCGCAGAGGTG	1080
Db	341	CysAlaAspSerAlaThrSerGlnAlaAlaGlyIleIyGlnArgProValHisProAla----	358
QY	1081	CCGAGCGTGGCCAGCGTGCAGCGCCTGCTGTTAAGAGACAGCCCGCGTGTACAGCGCTC	1140
Db	359	---SerValProSerValGlnProSerIleuIyGlnAspSerProIleuValGlnAlaIleu	377
QY	1141	AGTGGGCTGCACCTGTCCGACAGAGGCTGGAAACAGCAATTAAGCTCAAGCTTCTCTTC	1200
Db	378	SerGlyIleuHisIleuSerAlaAspArgIleuGlnAspSerIyAsnIySleuIyArgSerPhe	397

OY	1201	TCTGTGAAATCAATAACAGTATTCAATTTCAGCAGACAGAGAGATCTTTAATGGCTTC	1260
Dd	398	SerLeuAspIleLysSerValSerTyrlserAlaSerMetAlaAlaSerLeuHisgLyPhe	417
OY	1261	TCTCATCAAGAAGATGCTTTGGAAATATCAAAAACCTTCCACTATCTGGATGGAGCCAAC	1320
Dd	418	SerSerSerGlnAspIleAlaLeuGlylurTyrllysProSerThrThrLeuAspGlyThrAsn	437
OY	1321	AACSTATGCGACTTTCTCCCTGTGTATAGAACTATTCGAGCAGACTCCCGAAACCAGTCT	1380
Dd	438	LysLeuLysGlnPheSerProValGlnGluLeuSerGlnGlnInthrProGlnInthrSerPro	457
OY	1381	GATAGAGAGAGAAAGCAGACATCCCGCAAAGAGCTGGACAACCGGCAGGCTTTCAGACAGCAG	1440
Dd	458	AspIylsGlnGlnAlaSerIleProIylalysLeuGlnInthrAlaArgProSerAspSerGln	477
OY	1441	AGCAACGATTTGCATTTCGATCGACAAACAGCAGCAGTGGCACCGGCCAGAGAGTCCCTTTTA	1500
Dd	478	SerTylsArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu	497
OY	1501	TCTCCACATGCATTCGAAGTGGAGGGGTGGAGACAAATTACACACACAGCTTCTTTTCGCG	1560
Dd	498	SerProLeuHisArgSerGlySerValGlnAspAsnTrpHisThrSerPheLeuPheGly	517
OY	1561	CTTTCCACAGCAGCAGCAGCAGCTCACGAAGCTGTGGCTGGAGCCTTAAAGGCTGTGGCAC	1620
Dd	518	LeuSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuGlyLeuLysGlyIlePheHis	537
OY	1621	TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACCGAGCAGCTGTATTTTGGC	1680
Dd	538	SerAspIleLeuAlaProGlnInthrSerThrProSerLeuThrSerSerTrpIyrPheAla	557
OY	1681	ACAGAGTCCCTCAGACTTCTACTGCGGCTTAGCCATACGAGAGGACAGTGGCCAGTTACTCT	1740
Dd	558	ThrlusSerSerHisPheTyrlSerAlaSerAlaIleTyrlGlyIleSerAlaSerTyrlSer	577
OY	1741	GCTTACAGCTGCAGCCAGCTGCCACTTGTGGAGACCAAGTCTAATCTGTGTGGCAGAGCGG	1800
Dd	578	AlaIylrSerTyrlSerSerGlnLeuProThrCylsGlyAspGlnValTylSerValAlaArgIylArg	597
OY	1801	CAGAAAGCCAGTGCACAGAGCTGACTGCGGGCGGAGCTGGCATGAAGAGACCCCCTTTGAA	1860
Dd	598	GlnIylsProSerAspIleAlaSerSerAlaArgSerTrpHisGlnGlyIleSerProPheGln	617
OY	1861	AAGCAGTTTAAAGCGAAGAGCTGCCAAATGGAATTTGGAGAGAGACATATGTCAGAGAAC	1920
Dd	618	LysGlnIlePheLysArgIylSerCylsGlnMetGlnPheGlyGlnSerIleMetSerIleAsn	637
OY	1921	AGCTCAGGGAGAGAGCTGGGAAAGTGGCAGCTAGCTTACGTTTTGGGAGCAGATGGAA	1980
Dd	638	ArgSerArgGlnGlnLeuGlyLysValGlySerGlnInsSerPheSerCylsSerMetGln	657
OY	1981	ATCATTTGAGGTTCTCC	1995
Dd	658	IleIleIleIleValSer	662

Search completed: June 21, 2004, 13:05:33
Job time : 143.902 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:26 ; Search time 25.4447 Seconds

(without alignments)
8095.510 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 3615
Sequence: 1 atggcccatgagatgattg.....tggaaatcatgaggtctcc 1995

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=x1p
-Q=/cgn2_1/USPTO.spool_p/US10029345/runat_21062004_122816_4233/app_query.fasta_1.3278
-DB=Issued_Patents_AA -OPMT=fastaan -SUFFIX=rai -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTWRT=pct -NOM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USRT=US10029345_@C@G_1_1_45 @runat_21062004_122816_4233 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -MAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3406	94.2	665	US-09-816-494-2	Sequence 2, App1
2	654.5	18.1	170	US-09-544-716-14	Sequence 14, App1
3	654.5	18.1	170	US-09-557-921-15	Sequence 15, App1
4	654.5	18.1	170	US-09-564-357-17	Sequence 17, App1
5	654.5	18.1	170	US-09-619-380-16	Sequence 16, App1
6	469	13.0	482	US-09-557-921-2	Sequence 2, App1
7	455	12.6	394	US-09-702-705-805	Sequence 805, App
8	455	12.6	394	US-09-736-457-805	Sequence 805, App
9	455	12.6	394	US-09-614-1248-805	Sequence 805, App
10	455	12.6	394	US-09-671-325-805	Sequence 805, App
11	455	12.6	394	US-09-589-184-805	Sequence 805, App
12	447	12.4	394	US-08-530-290-23	Sequence 23, App1

13	447	12.4	394	US-09-702-705-827	Sequence 827, App
14	447	12.4	394	US-09-736-457-827	Sequence 827, App
15	447	12.4	394	US-09-614-1248-827	Sequence 827, App
16	447	12.4	394	US-09-671-325-827	Sequence 827, App
17	447	12.4	394	US-09-589-184-827	Sequence 827, App
18	439.5	12.2	395	US-08-990-379-5	Sequence 5, App1
19	433	12.0	367	US-08-990-379-6	Sequence 6, App1
20	425	11.8	314	US-09-371-6718-11	Sequence 11, App1
21	420	11.6	367	US-08-530-290-24	Sequence 24, App1
22	409.5	11.3	313	US-08-990-379-7	Sequence 7, App1
23	409	11.3	314	US-09-164-193-22	Sequence 22, App1
24	409	11.3	314	US-09-221-448A-22	Sequence 22, App1
25	393.5	10.9	393	US-08-990-379-4	Sequence 4, App1
26	377	10.4	302	US-09-702-705-806	Sequence 806, App
27	377	10.4	302	US-09-736-457-806	Sequence 806, App
28	377	10.4	302	US-09-614-1248-806	Sequence 806, App
29	377	10.4	302	US-09-671-325-806	Sequence 806, App
30	377	10.4	302	US-09-589-184-806	Sequence 806, App
31	366	10.1	397	US-08-990-379-8	Sequence 8, App1
32	346.5	9.6	168	US-09-544-716-13	Sequence 13, App1
33	346.5	9.6	168	US-09-557-921-13	Sequence 13, App1
34	346.5	9.6	168	US-09-564-357-16	Sequence 16, App1
35	346.5	9.6	168	US-09-619-380-15	Sequence 15, App1
36	338.5	9.4	170	US-09-544-716-12	Sequence 12, App1
37	338.5	9.4	170	US-09-557-921-12	Sequence 12, App1
38	338.5	9.4	170	US-09-564-357-15	Sequence 15, App1
39	338.5	9.4	170	US-09-619-380-14	Sequence 14, App1
40	319	8.8	169	US-09-544-716-16	Sequence 16, App1
41	319	8.8	169	US-09-557-921-17	Sequence 17, App1
42	319	8.8	169	US-09-564-357-19	Sequence 19, App1
43	319	8.8	169	US-09-619-380-18	Sequence 18, App1
44	312	8.6	168	US-09-544-716-15	Sequence 15, App1
45	312	8.6	168	US-09-557-921-16	Sequence 16, App1

ALIGNMENTS

RESULT 1
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816, 494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-816-494-2

Alignment Scores:
Pred. No.: 7.74e-300
Score: 3406.00
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 94.22%
DB: 4
Gaps: 0

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-816-494-2 (1-665)

QY 1 ATGGCCCATGAGATGATTGGAAGTCAATTTGTAACGAGAGCTTGCTGCTGCGA 60
Db 1 MetAlhIeGlWetIleGlyThrGlnIleValThhGluArgIleValAlaIleuLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTGTGCTAATTGATGATGCGGCACTTTGTGAATACATATCC 120

D	b		381 SerGlyLeuHisIleSerSerAlaAspArgLeuGlnAspSerIleuLysArgSerPhe 400
O	y	1201	TCTGTGAATTAACAATCAGTTTCATTATTCAGCCACGACTGGCACATCTTACATGGCTTC 1260
D	b	401	SerIleuAspIleLysSerValSerTyrrSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
O	y	1261	TCCCATCAAGAAGATCGTTTGAAATCTAACAAACCTTGCATACTTGGATGGAGCAAC 1320
D	b	421	SerSerSerGluAspAlaLeuGlnTyrrTyrrLysProSerThrThreuleuAspLyrThrAsn 440
O	y	1321	AAGGTATGCCAGTTCCTCCCCTGTTCCAGAACTATCGAGACAGACATCCCCGAACAGTCCT 1380
D	b	441	LysIleuCysAspIlePheSerProValGlnGlnIleuSerGlnGlnThrProGlnThrSerPro 460
O	y	1381	GATTAAGAGAGAAACCAGCATCCCCCAAAGAACGTGCAGACCCGCCAGGCTTTCAGACGCCAG 1440
D	b	461	AspLysGlnGlnIleuAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
O	y	1441	AGCAGACCGATTCGATTCGGTCAAGAACCGACGACGATGGCAGCCGACAGAGTCCCTTTA 1500
D	b	481	SerLysArgLeuHisIleSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
O	y	1501	TCTCCATGTCATTCGAAGTGGAGAGCGTGGAGAGCAATTAACACACACAGCTTCTTTTCGGC 1560
D	b	501	SerProLeuHisIleArgSerGlySerValGlnAspAsnTrpHisThrSerPheIleuPheGly 520
O	y	1561	CTTTCACACGACCGACGACGACCTTACGAAGTCTGTGGCCTGGGCTTAAGGAGCTGGAC 1620
D	b	521	LeuSerThrIserGlnGlnHisIleThrLysSerAlaGlyLeuGlyLeuLysGlyTrpHis 540
O	y	1621	TGGGATATCTTGGCCCCCGACGACCTTACCCCTCCCTGACACGACGCGATATTTGGC 1680
D	b	541	SerAspIleLeuAlaProGlnThrSerThrProSerIleThrSerSerTrpTyrrPheAla 560
O	y	1681	ACAAGAGCTCCACACTTCTACTCTGCTCGACGACCTTACGAGGAGCGAGTCCAGTTACTCT 1740
D	b	561	ThrGlnSerSerHisPheTyrrSerAlaIleTyrrGlyGlySerAlaSerTyrrSer 580
O	y	1741	GCTTACAGCGTCGACGACGACGCTGCCCACTTGGCAGAGCAAGTCTATTCTGTGAGGAGCGG 1800
D	b	581	AlaTyrrSerCysSerGlnIleuProThrTyrrGlyAspGlnValTyrrSerValArgArgArg 600
O	y	1801	CAGAACCCAAAGTAGACAGACTGACTCGCGCGAGCGTGGCATGAAGAGAGGCCCTTTGAA 1860
D	b	601	GlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGlnGlnIleuSerProPheGln 620
O	y	1861	AAGCAGTTTAAACCGAGAAAGCTGCAATATGCAATTTGAGAGAGCAATCATGTACAGAAAC 1920
D	b	621	LysGlnPheLysArgLysSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
O	y	1921	AGGTACACGGAGAAAGCTGGGAGAAAGTGGGACGTCACTTTCCTTGGCGACATGGAA 1980
D	b	641	ArgSerArgGlnGlnIleuLeuGlyLysValGlySerGlnSerPheSerGlySerMetGln 660
O	y	1981	ATCATTTAGAGTCTCC 1995
D	b	661	IleIleGlnValSer 665
 RESULT 2 US-09-544-716-14 Sequence 14, Application US/09544716 Patent No. 6492157 GENERAL INFORMATION: APPLICANT: Lucche, Ralf M. APPLICANT: Wei, Bo TITLE OF INVENTION: DSP-9 DUAL-SPECIFICITY PHOSPHATASE FILE REFERENCE: 200125_415 CURRENT APPLICATION NUMBER: US/09/544_716 CURRENT FILING DATE: 2000-04-10 NUMBER OF SEQ ID NOS: 20 SOFTWARE: FastSeq for Windows Version 4.0 SEQ ID NO 14 LENGTH: 170			

TYPE: PRT
ORGANISM: Homo sapiens
US-09-544-716-14

Alignment Scores:

Pred. No.: 6,75e-51 Length: 170
Score: 654.50 Matches: 121
Percent Similarity: 88.62% Conservative: 27
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 18.11% Indels: 1
DB: 4 Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-544-716-14 (1-170)

QY 400 GGCCTCTGTGAGAGAAA--TCACCTTATGTCCTTACCTGCAATTTCTCAGCCTTGCTTA 456
DB 1 GtLeuCyGsluGluYlsProAlaAlaLeuLeuProMetSerLeuSerGlnProCySleu 20
QY 457 CCTGTGGCCAAACATTTGGGCCAACCCGAATCTTCCCAATCTTATCTTATGCTGGCCAGCGA 516
DB 21 ProValProSerValGtLeuThrArgGtLeuProHtLeuProHtLeuGtLeuGtLeuGtLeu 40
QY 517 GATGTCCTCAACAGAGAGCTGATACAGAGATGGATGGATTGTTATGTTAAATGGCAGC 576
DB 41 AspValLeuAsnLysAspLeuMetThrGlnAsnGtYlLeSerYrValLeuAsnAlaSer 60
QY 577 TATACCTGTCAAAGCCTGACTTATCCCGAGTCTCATTTCCGTGCTGCTGAT 636
DB 61 AsnSerCyProLysProAspPheIleCyGsluSerArgPheMetArgValProIleAsn 80
QY 637 GACAGCTTTGTGAGAAAATTTTGGCGTGTGAGCAAAACAGATGTTCAATGAGAA 696
DB 81 AspAsnTyrCyGsluYlsLeuLeuProTrpLeuAspLysSerIleGluPheIleAspLys 100
QY 697 GCAAAAGCCTCCAAATGATGTCTTACGACATGTTTACGCGATCTCCCGCTCCGCC 756
DB 101 AlAluYlsLeuSerSerCySglnValIleValHtCySleuAlaGtYlLeSerArgSerAla 120
QY 757 ACCATGCTATCGCTTACATCATGAGAGAGATGATGCTTATGATGAGCTTACAGA 816
DB 121 ThrIleAlaIleAlaTyrIleMetYrThrMetGtYlMetSerSerAspAlaTyrArg 140
QY 817 TTTGTGAGAAAAGAAAGAACTTACTATATCTCCAAATCTTATTTCTGGGCCAACTCTG 876
DB 141 PheValLysAspArgArgProSerIleSerProAsnPheAsnPheLeuGtYlGlnLeu 160
QY 877 GACTATGAGAGAAGATTAG 897
DB 161 GluTyrGluArgThrLeuLys 167

RESULT 3

US-09-557-921-15
Sequence 15, Application US/09557921
Patent No. 6551810
GENERAL INFORMATION:
APPLICANT: Lucbe, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125, 416
CURRENT APPLICATION NUMBER: US/09/557, 921
CURRENT FILING DATE: 2000-04-20
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 170
TYPE: PRT
ORGANISM: Homo sapiens
US-09-557-921-15

Alignment Scores:
Pred. No.: 6,75e-51 Length: 170
Score: 654.50 Matches: 121
Percent Similarity: 88.62% Mismatches: 27

Best Local Similarity: 72.46% Mismatches: 18
Query Match: 18.11% Indels: 1
DB: 4 Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-557-921-15 (1-170)

QY 400 GGCCTCTGTGAGAGAAA--TCACCTTATGTCCTTACCTGCAATTTCTCAGCCTTGCTTA 456
DB 1 GtLeuCyGsluGluYlsProAlaAlaLeuLeuProMetSerLeuSerGlnProCySleu 20
QY 457 CCTGTGGCCAAACATTTGGGCCAACCCGAATCTTCCCAATCTTATCTTATGCTGGCCAGCGA 516
DB 21 ProValProSerValGtLeuThrArgGtLeuProHtLeuProHtLeuGtLeuGtLeuGtLeu 40
QY 517 GATGTCCTCAACAGAGAGCTGATACAGAGATGGATGGATTGTTATGTTAAATGGCAGC 576
DB 41 AspValLeuAsnLysAspLeuMetThrGlnAsnGtYlLeSerYrValLeuAsnAlaSer 60
QY 577 TATACCTGTCAAAGCCTGACTTATCCCGAGTCTCATTTCCGTGCTGCTGAT 636
DB 61 AsnSerCyProLysProAspPheIleCyGsluSerArgPheMetArgValProIleAsn 80
QY 637 GACAGCTTTGTGAGAAAATTTTGGCGTGTGAGCAAAACAGATGTTCAATGAGAA 696
DB 81 AspAsnTyrCyGsluYlsLeuLeuProTrpLeuAspLysSerIleGluPheIleAspLys 100
QY 697 GCAAAAGCCTCCAAATGATGTCTTACGACATGTTTACGCGATCTCCCGCTCCGCC 756
DB 101 AlAluYlsLeuSerSerCySglnValIleValHtCySleuAlaGtYlLeSerArgSerAla 120
QY 757 ACCATGCTATCGCTTACATCATGAGAGAGATGATGCTTATGATGAGCTTACAGA 816
DB 121 ThrIleAlaIleAlaTyrIleMetYrThrMetGtYlMetSerSerAspAlaTyrArg 140
QY 817 TTTGTGAGAAAAGAAAGAACTTACTATATCTCCAAATCTTATTTCTGGGCCAACTCTG 876
DB 141 PheValLysAspArgArgProSerIleSerProAsnPheAsnPheLeuGtYlGlnLeu 160
QY 877 GACTATGAGAGAAGATTAG 897
DB 161 GluTyrGluArgThrLeuLys 167

RESULT 4

US-09-564-357-17
Sequence 17, Application US/09564357
Patent No. 6645753
GENERAL INFORMATION:
APPLICANT: Lucbe, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-5 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125, 413
CURRENT APPLICATION NUMBER: US/09/564, 357
CURRENT FILING DATE: 2000-04-24
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 170
TYPE: PRT
ORGANISM: Homo sapiens
US-09-564-357-17

Alignment Scores:
Pred. No.: 6,75e-51 Length: 170
Score: 654.50 Matches: 121
Percent Similarity: 88.62% Conservative: 27
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 18.11% Indels: 1
DB: 4 Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-564-357-17 (1-170)
QY 400 GGCCTCTGTGAGAGAAA--TCACCTTATGTCCTTACCTGCAATTTCTCAGCCTTGCTTA 456
DB 1 GtLeuCyGsluGluYlsProAlaAlaLeuLeuProMetSerLeuSerGlnProCySleu 20

```

Db      1 GlyLeuCysGluGlyIysProAlaLeuLeuProMetSerLeuSerGlnProCysLeu 20
QY      457 CCTGTGCGCAACATTGGGCCAACCCGAATTTCTTCCAAATCTTTATCTTGGCTGCCACGCA 516
Db      21 ProValProSerValGlyLeuThrArgIleLeuProHisLeuGlySerGlnIys 40
QY      517 GATGTCTCAACAAGAGCTGATACAGCAAGATGGAGATGGTTATGTGTTAAATGCGCAGC 576
Db      41 AspValLeuAsnIysAspLeuMetThrGlnAsnGlyIleSerIleValLeuAsnAlaSer 60
QY      577 TATACCTGCGCAAGCTGACTTATATCCCGAGTCTCATTTCCGCGGTGGCTGGAT 636
Db      61 AsnSerCysProIysProAspPheIleCysGlnSerArgPheMetArgValProIleAsn 80
QY      637 GACAGCTTTGTGAGAAATTTTGGCGGTGGAGCAATTCAGTATGATGAAGCTTACAGA 696
Db      81 AspAsnIleCysGlnIysLeuLeuProIlePheAspIysSerIleGlnPheIleAspIys 100
QY      697 GCAAAAGCTCCAAATGATGTGTTCTAGTGCACGTGTTAGCTGGAGATCCCGCTCCGCC 756
Db      101 AlalysLeuSerSerCysGlnValIleValHisCysLeuAlaGlyIleSerArgSerAla 120
QY      757 ACCATGCTATGCGCTCATCATGAGAGAGAGATGATGCTTTAGATGAAGCTTACAGA 816
Db      121 ThrIleAlaIleAlaIleTyrIleMetIysThrMetGlyMetSerSerAspAlaIleTyrArg 140
QY      817 TTGTGAAAGAAAAGAAAGCTACTATATCTTCCAAATTTCTTGGCGCAACTCTG 876
Db      141 PheValIysAspArgPheProSerIleSerProAsnPheAsnPheLeuGlyGlnLeuLeu 160
QY      877 GACTATGAGAAAGATTAAG 897
Db      161 GlutyrGlnArgThrLeuIys 167

```

RESULT 5
US-09-619-380-16
Sequence 16, Application US/09619380
Patent No. 6649391
GENERAL INFORMATION:
APPLICANT: Lucbe, Ralf M.
TITLE OF INVENTION: DSP-11 DUAL SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.418
CURRENT APPLICATION NUMBER: US/09/619,380
CURRENT FILING DATE: 2000-07-19
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 170
TYPE: PRN
ORGANISM: Homo sapiens
US-09-619-380-16

Alignment Scores:
Pred. No.: 6,75e-51
Score: 654.50
Percent Similarity: 88.62%
Best Local Similarity: 72.46%
Query Match: 18.11%
Matches: 121
Conservative: 27
Mismatch: 18
Indels: 1
Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-619-380-16 (1-170)

```

QY      400 GGCCCTGTGAAAGAAA---TCACCTATGCTCCCTACGCTTCTTCTGACCTTGTCTTA 456
Db      1 GlyLeuCysGluGlyIysProAlaLeuLeuProMetSerLeuSerGlnProCysLeu 20
QY      457 CCTGTGCGCAACATTGGGCCAACCCGAATTTCTTCCAAATCTTTATCTTGGCTGCCACGCA 516
Db      21 ProValProSerValGlyLeuThrArgIleLeuProHisLeuGlySerGlnIys 40
QY      517 GATGTCTCAACAAGAGCTGATACAGCAAGATGGAGATGGTTATGTGTTAAATGCGCAGC 576

```

```

Db      41 AspValLeuAsnIysAspLeuMetThrGlnAsnGlyIleSerIleValLeuAsnAlaSer 60
QY      577 TATACCTGCGCAAGCTGACTTATATCCCGAGTCTCATTTCCGCGGTGGCTGGAT 636
Db      61 AsnSerCysProIysProAspPheIleCysGlnSerArgPheMetArgValProIleAsn 80
QY      637 GACAGCTTTGTGAGAAATTTTGGCGGTGGAGCAATTCAGTATGATGAAGCTTACAGA 696
Db      81 AspAsnIleCysGlnIysLeuLeuProIlePheAspIysSerIleGlnPheIleAspIys 100
QY      697 GCAAAAGCTCCAAATGATGTGTTCTAGTGCACGTGTTAGCTGGAGATCCCGCTCCGCC 756
Db      101 AlalysLeuSerSerCysGlnValIleValHisCysLeuAlaGlyIleSerArgSerAla 120
QY      757 ACCATGCTATGCGCTCATCATGAGAGAGATGATGCTTTAGATGAAGCTTACAGA 816
Db      121 ThrIleAlaIleAlaIleTyrIleMetIysThrMetGlyMetSerSerAspAlaIleTyrArg 140
QY      817 TTGTGAAAGAAAAGAAAGCTACTATATCTTCCAAATTTCTTGGCGCAACTCTG 876
Db      141 PheValIysAspArgPheProSerIleSerProAsnPheAsnPheLeuGlyGlnLeuLeu 160
QY      877 GACTATGAGAAAGATTAAG 897
Db      161 GlutyrGlnArgThrLeuIys 167

```

RESULT 6
US-09-557-921-2
Sequence 2, Application US/09557921
Patent No. 6551810
GENERAL INFORMATION:
APPLICANT: Lucbe, Ralf M.
TITLE OF INVENTION: DSP-10 DUAL SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.416
CURRENT APPLICATION NUMBER: US/09/557,921
CURRENT FILING DATE: 2000-04-20
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 482
TYPE: PRN
ORGANISM: Homo sapiens
US-09-557-921-2

Alignment Scores:
Pred. No.: 7.49e-34
Score: 469.00
Percent Similarity: 57.72%
Best Local Similarity: 35.91%
Query Match: 12.97%
Matches: 107
Conservative: 65
Mismatch: 96
Indels: 30
Gaps: 7

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-557-921-2 (1-482)

```

QY      79 CTGCTAATGATAGCGCGCATTTGTGGAATACATACATCCCAATTTGGAAGCATT 138
Db      173 ValIleIleAspCysArgProPheMetClnIysAsnIysSerHisIleGlnIleAlaVal 192
QY      139 AATATCAACTGCTCC---AAGCTATAGACGAGAGGTTGACAGCAAGAGGTTAATT 195
Db      193 HisIleAsnCysAlaAspIysIleSerArgArgIleGlnIleGlnIleThrVal 212
QY      196 ACAGAGCTATCCAG---CAATTCAGCAAGCAATTAAGATTGACATGATTCAGTCAGAG 252
Db      213 LeuAspLeuIleSerCysArgGluGlyIysAspSerPheIysArgIlePheSerIysGln 232
QY      253 GTTGAAGTTATACATCAAACTCCCAAGAGTGGCTCTCTCTTCAAGACTGTTTCTC 312
Db      233 IleIleValTyrAspGlnAsnThrAsnGlnProSerArgValMetProSerGlnProLeu 252
QY      313 ACTGACTCTGCGGTAACTGAGAGAGAGCTTCAACTGTTTCACTGCTTGCAGTGCGG 372

```

```

Db      253 HieilValleuInserLeuLyArgGlUgLyseglProleuValleuLyseglY 272
Qy      373 TTGGCTGAGTCTCTGCTGTTTCCCTGGGCTCTGTGACAGAAAA----- 417
Db      273 LeuSerSerHeuLyseglInseNH1sglUAsenLeuCyseApasenSerleuInseglU 292
Qy      418 -----TCCACTAGCCCTTACCTGATTTCT 444
Db      293 CyseArgGlUValGelYglYalaseRAlAlaseSerleuProglInProle--- 311
Qy      445 CAGCTTGCTTACTCT--GTGGCAACATTGGGCCAACCCGAATTTCTCCCACTTTAT 501
Db      312 ---ProthThrProAspRleglUAsenAlglUenThrProileuProPheleuPhe 330
Qy      502 CTGGGCTGCCAGAGATGTCTTCACACAGAGAGCTATACAGCAAGATGGATGGTTAT 561
Db      331 LeuGlYAsnGlUAsnGlUAsnRAlAsnRleuAspThrMetGlInArgLeuAsnIlleGlY 350
Qy      562 GTGTAAATGSCAG-----TATACCTGTCAAGAGCTGACTTATCCCC 606
Db      351 ValIleAsnValInThrThIsleuProleuThyThIsTglUlyseglUleuPhe----- 368
Qy      607 GAGTCTATTCTCGCTGTGCTGTGATGACAGCTTTGTGAGAAATTTTGCCTGTG 666
Db      369 -----AsnTyrLyAspArgLeuProAlaThrAspSerAsnLyseglInseNH1ArgGlN 386
Qy      667 TTGGACAAATCACTAGTATTTCATTGAGAAAGCAAAAGCCCTCAATGATGTTCTATG 726
Db      387 PheGlUglUAlaPheglUAsnRleuInseglUAlaNH1seglUyAsglYlyseglUleuIlle 406
Qy      727 CACTGTATGAGGATGATCCCGCTCGCACCAAGCTGTGCTATGCAATGATGAGAGG 786
Db      407 HisCyseglInAlaRleYAlaSerArgSerAlaThrIleAlaIleAlaTyrleuMetLysNH1 426
Qy      787 ATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTTATATCT 846
Db      427 ThrArgMetThrMetThrAspAlaTyrLysePheVallyseglYAspArgProIleIleSer 446
Qy      847 CCAACTTCATTTTCTGGGCGCAATCTCTGAGCTATGAGAAAGATTAAGAAC 900
Db      447 ProAsnLeuAsnRheMetGlyInLeuInseglUAsnRleuInseglUAsnRleuAsn 464

RESULT 7
US-09-702-705-805
; Sequence 805, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvik, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-705-805

Alignment Scores:
Pred. No.: 1,25e-32 Length: 394
Score: 455.00 Matches: 122
Percent Similarity: 49.87% Conservative: 75

```

```

Best Local Similarity: 30.89% Mismatches: 143
Query Match: 12.59% Indels: 55
Db: 4 Gaps: 13

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-702-705-805 (1-394)

Qy      45 GGTGCTCTGCTGAGAAAGTGAACGGA-----AAAGT 77
Db      26 GlylYAlaGlYglY-SerGlYserNH1seglYThrleuGlYleuProSerGlYglYyScy 45
Qy      78 GCTGTAATGATAGCGGCCATTTGTGAAATACATTAATCCCAATTTGGAAGCAT 137
Db      45 sleuLeuLeuAspCyAspArgProPheleuAlaNH1SerAlaGlYTyrIleleuGlYserVa 65
Qy      138 TAATATCACTGTCTCCAGCTTATGAAAGCAAGTTGCAACAGACAAAGTGAATAC 197
Db      65 lAsnValArgCyAsnThrIleValArgArgArg--AlaLyseglYserAlaSerleuGl 84
Qy      198 AGAGCTCATC-----CAGATTCAAGCAAACTAAGTTGACATTGACATGCAGAA 251
Db      84 uGlInIleuProAlaGlUglUglUAlaRAlaRAlaRleuArgSerGlYleuTyrSerAl 104
Qy      252 GGTGTAGATTAGATCAAAAGCTCCCAAGATGTGCTCTCTTCAAGCTGTTTCT 311
Db      104 aValIleValTyrAspRglUAspSerProArgAlaInseRleuArgGlUAspSerThrVa 124
Qy      312 CACTGATCTTGTGGTAAATGGAAGAGCTTC-----AACTGTTCACCTGCTTGC 365
Db      124 lSerleuValAlaGlAlaIleuArgArgAsnAlaGlUArgThrAspIleCyseleuLeu 144
Qy      366 AGGTGGTTTGTGCTGATGTTCTGTGTTTCCCTGAGCTGTGAAAGAAATCCACTG 425
Db      144 seglYAlYtYglUArgPheSerSerleuTyrProglUAspCyseSerleYThrYAlaIle 164
Qy      426 A-----GTCCCTACTGCATTTCTAGCT-----TGCTT 455
Db      164 uAlaAlaIleProProProAlaProProSerAlaThrInglUProleuAspRleuAspCyse 184
Qy      456 ACCTGTGCC-----AACATGGGCCAACCCGAATTTCTCCCAATCTTGA 500
Db      184 rSerCyseglYThrProleuNH1AspRlInseglUProValglUAlleleuProHeuLeu 204
Qy      501 TCTTGGCTGACAGAGATGTCTTCAAGAGAGCTGATACAGAGATGATGGTTA 560
Db      204 rleuGlYserAlaTyrNH1AlaAlaArgArgMetleuAspRleuGlYleThrAl 224
Qy      561 TGTGTAAATGSCAGCTATACCTGTCCAAAGCTGATTAATCCCGAGCTCATTTCT 620
Db      224 aleuLeuAsnValSerSerAspCyseProAsn--HisPheglUglYNH1TyrGlInTyrLy 243
Qy      621 GCCTGTGCTGTGATGACAGCTTTTGTGAAAAATTTGCGCTGTGACAAATCACT 680
Db      243 eCyAlleProValRglUAspAsnNH1eYAlaAspIleSerSerTyrPheMetGlUAlaIl 263
Qy      681 AGATTCATTGAGAAAGCAAAAGCTCCCAATGATGTTGTATGACATGTTAGCTGG 740
Db      263 eglUtyrIleAspAlaVallyseNH1AspCyseArgglYArgValleuValHisCyseGlNH1agl 283
Qy      741 GATCTCCGCTCGCACCATGCTATCCCTTACATCAAGAGAGATGAGACATGTCTTT 800
Db      283 yIleSerArgSerAlaThrIleCyseuAlaTyrleuMetMetlyLyAlaRAlaRgl 303
Qy      801 AGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTATATCTTCCCAATCTTCAATT 860
Db      303 uGlUglUAlaPheglUAsnRleuVallyseglInArgArgSerIleIleSerProAsnRheSerP 323
Qy      861 TCTGGGCAAGCTCTGAGATGAGAAAGATTAAGACAGACAGACTGAGAGAGGCC 920
Db      323 emetGlYInleuLeuInPheglUserInValleuAlaThrNH1SerCyAlaAlaIleAla 343
Qy      921 AAAGACAAATCTCAAGCTGTGACCTGAGAGAGCAAAATGAACCTGTGCTGTCTC 980
Db      343 aAlaSer-----ProSerGlYProleu----- 350

```

QY 981 AGAGGGTGGACAGAAAAAGCAGACGCCCTCTGATCCACCGTGTGCCGACTCTGTCACTC 1040
 Db 351 -----GlyIuArgGlyLysThrProAlaThrPro-----ThrSe 362
 QY 1041 AGAGGACAGACAGACAAAGCCCGGTGCATCCGCCAGCGTGCACCGCGTACCCAGCGATCA 1100
 Db 362 rcGlnPheValPheSerPheProVal-----SerValGlyValLHisSerAlaProSerSerLe 381
 QY 1101 GCCGTCGCTGTTAAGAGACAGACGCCCGCTGTGTACAGCGTCACGT 1143
 Db 381 uProTyrLeu-----HisSerProIleThrThrSerProSer 393

RESULT 8

```

US-09-736-457-805
? Sequence 805, Application US/09736457
? Patent No. 6509448
? GENERAL INFORMATION:
? APPLICANT: Wang, Tonglong
? APPLICANT: Bangur, Chaltanya S.
? APPLICANT: Lodes, Michael A.
? APPLICANT: Fanger, Gary
? APPLICANT: Vedvick, Tom
? APPLICANT: Carter, Dartrick
? APPLICANT: Retter, Marc
? APPLICANT: Mannion, Jane
? APPLICANT: Fan, Liqun
? APPLICANT: Wang, Aijun
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
? FILE REFERENCE: 210121.478C15
? CURRENT APPLICATION NUMBER: US/09/736,457
? CURRENT FILING DATE: 2000-12-13
? NUMBER OF SEQ ID NOS: 1864
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 805
? LENGTH: 394
? TYPE: PR1
? ORGANISM: Homo sapiens
US-09-736-457-805

```

Alignment Scores:

File: NO.:	1.25e-32	Length:	3
Score:	455.00	Matches:	75
Percent Similarity:	49.87%	Conservative:	12
Best Local Similarity:	30.89%	Mismatches:	14
Query Match:	12.59%	Gaps:	55
DB:	4	Indels:	13

...CONF_530_2532 (1-1995) X US-09-736-457-805 (1-394)

```

47 45 GGNAGCTCTGCTGGAAAGTGGAAAGAA-----AAAGT 77
Db 26 GlyGlyAlaGlyGly-SerGlySerHisGlyThrLeuGlyLeuProSerGlyGlyLeuGlyC 45
Qy 78 GCTGCTAATTGATAGCCGCGCCATTGTGTGAATACATATCCCATTTTGGAAAGCC 13
Db 45 sLeuLeuLeuaspPysArgProPheLeuAlaHisSerAlaGlyTyrIleLeuGlySerVal 65
Qy 138 TAATATCAACTGCTGCCAAGCTTATGAAGGAAGCTTGCAACAGCAAGCAAGATTAAATAC 197
Db 65 IAsnValArgCysAsnThrIleValArgArgArg--AlaYsgIserValSerLeuG 84
Qy 198 AGAGCTCATC-----CAGCATTCAGCGAAACATATAGGTTGACATTTAGTTCAGTCA 251
Db 84 uGlnIleLeuProAlaGlnGluGluValArgAlaArgLeuArgSerGlyLeuTyrSerAl 104
Qy 252 GGTGTAGTATTACGATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCAAGACTGTTTATC 311
Db 104 aValIleValTyrAspGluArgSerProArgAlaGlnSerLeuArgGluAluAspSerThrVal 124
Qy 312 CACTGACTCTTGGGTAACTGGAGAAGAGCTTC-----AATCTGTTCACCTGTTTGC 365
      ::::::::::::::: ||| ::::::::::: AATCTGTTCACCTGTTTGC 365

```

Db -124 lserleuValValGlnAlaIeAluArgAlaAsnAlaGluArgThrAsp1IeCysLeuLeu 144

QY 366 AGGTGGGTTCGTAAGTTCCTGCTGTTGTCCTCGGGCTCTGTGAAGAAATTCACCTT 425

Db 144 sGlyIYrYrGluArgPheSerSerGluTYrProGluPheCysSerIYrThrIYsAlaLe 164

QY 426 A-----GTCCCACTCGATTTCTCAAGCT-----TGCTT 455

Db 164 uAlaAlaIaIeProProProValProProSerAlaThhGluProLeuAlaAspLeuAspCysSe 184

QY 456 ACCTGTGGC-----AACATYTGCCCAACCCGAATTTCTTCCCAATCTTGA 500

Db 184 rSerYsGluYrThrProLeuHIsaSpGlnGluYrProValGluIleuProPheLeu 204

QY 501 TCTTGGTCGCCAGCAGATGCTCCAAAGAGAGCTGATACAGCAAGATGGGATTTGGTTA 560

Db 204 rIleuGlySerAlaYrHIsAlaIaIaArgAlaGAspMetLeuAlaAlaIeGluIleThraI 224

QY 561 TGTGTAAATGCCAGGCAATACCTGTCCAAAGCCTGATTTATCCCGAGTGCATTTCTT 620

Db 224 aIeuAlaAsnValIseSerAspCysProAsn--HisPheGlnGlyHIsYrGlnTYrIYr 243

QY 621 GCGTGTCCCTGTGAATGACAGCTTTTGTGAGAAATTTTGCCGTGTTGGACAAATAGT 680

Db 243 sCysIleProValGluAlaSpAsnHIsAlaIaAspIleSerSerTrpPheMetGluAlaI 263

QY 681 AGATTCATTTGAGAAAGCAAAAGCCTCCAAAGATGTGTTTAAGTCACTGTTTAGCTGG 740

Db 263 eGluTYrIleAlaPalValIYsAspCysAlaGlyIArgValIeValHIsCysGlnAlaG 283

QY 741 GATTCGCCGCTCCGACCACTTCGCTATTCGCTCATGATGAAGAGATGACATGACTTT 800

Db 283 yIleSerAlaSerAlaTrnIleCysLeuAlaTYrLeuMetMetIYsAlaYrValIArgIe 303

QY 801 AGATGAAGCTTACAGATTTTGAAGAAAGAAAGCCACTATATCTTCCAAATCTCAATT 860

Db 303 uGlnGlnAlaPheGlnPheValIYsGlnAlaGlySerIleIleSerProAsnPheSerPh 323

QY 861 TCTGGGCAACCTCCCTGAGCTATGAGAAAGAAATTAAAGCAAGACTGAGATCAAGGCG 920

Db 323 emetCylGlnIleuGlnPheGlnIleValIeAlaTrnSerCysAlaAlaGluAl 343

QY 921 AAAGGCAAACTCAAGCTGCTGACCTGAGAAAGCCAAATGAACCTGTCGCTGTCTC 980

Db 343 aAlaSer-----ProSerGlyProLeu----- 350

QY 981 AGAGGTGACAGAAAGAGCAGACGCCCTCAGTCAACCTGTGCGCATCTGTGTAACCTC 1040

Db 351 -----GlyGluArgGlyIYrThrProAlaTrnPro-----Thrse 362

QY 1041 AGAGGCGACAGCAAAAGCCCTGTGCATTCGCCACAGCTGCCCCAGCTGCCCCAGCGTGA 1100

Db 362 rGlnPheValPheSerPheProVal--SerValGlyValHIsSerAlaProSerSerIe 1101

QY 1101 GCGTGTGCTGTTAGAGACAGGCCGCTGTGTACAGCGCTCAGT 1143

Db 381 uProTYrIleu-----HisSerProIleTrnThrSerProSer 393

RESULT 9

US-09-614-124B-805
Sequence 805, Application US/09614124B
Patent No. 6630574
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaltanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvik, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

```
FILE REFERENCE: 210121.478C9
CURRENT APPLICATION NUMBER: US/09/614,124B
CURRENT FILING DATE: 2001-07-11
NUMBER OF SEQ ID NOS: 1668
SOFTWARE: FastSeq for Windows Version 3.0.
SEQ ID NO 805
LENGTH: 394
TYPE: PRT
ORGANISM: Homo sapiens
US-09-614-124B-805

Alignment Scores:
Pred. No.: 1,25e-32 Length: 394
Score: 455.00 Matches: 122
Percent Similarity: 49.87% Conservative: 75
Best Local Similarity: 30.89% Mismatches: 143
Query Match: 12.59% Indels: 55
DB: Gaps: 13

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-614-124B-805 (1-394)

QY 45 GGTGGCTGCTGCGAAAGTGAACGAA-----AAAGT 77
Db 26 GlyGlyAlaGly-SerGlySerHisGlyThrLeuGlyLeuProSerGlyGlySerCy 45
QY 78 GGTGTAATGTAAGCCGGCATTGTGGAATACATACATCCACATTTTGAAGCCAT 137
Db 45 GLeuLeuSerCyArgProPheLeuAlaHisSerAlaGlyTrpLeuGlySerVa 65
QY 138 TAATATCAATGCTGCACAGCTTATGAAAGCAGAGTTGCAACAGACAAAGTTTAATTAC 197
Db 65 IAsnValArgCyAsnThrIleValArgArg---AlaGlyGlySerValSerLeuG 84
QY 198 AGAGCTCANT-----CAGATTCAAGCAAACTAAGTTGACATTGACATGCAAGAA 251
Db 84 uGlnIleLeuProAlaGlyLeuGlyValArgAlaArgLeuArgSerGlyLeuTrpSerAl 104
QY 252 GGTGTAATGTAAGTGAACGCTCCCAAGATGTTGCTCTCTTCAGACTGTTTCT 311
Db 104 aValIleValTrpArgProIleArgSerProArgAlaGlySerLeuArgGlyAspSerThrVa 124
QY 312 CACTGTAATCTGGTGAACCTGGAAGAGACTTC-----ACCTGTTCACTGCTTGC 365
Db 124 IserIleValValGlnAlaLeuArgArgAsnAlaGlyArgThrAspIleCysLeuLeuGly 144
QY 366 AGGTGGTGTGCTGAGTTCTCTGCTGTTTCCCTGGCTCTGTAAGAAATTCACCTCT 425
Db 144 sGlyGlyTrpGlyArgPheSerSerGlyTrpProGluPheCysSerTrpThrValAla 164
QY 426 A-----GTCCCTACCTGCAATTCAGACT-----TGCTT 455
Db 164 uAlaAlaIleProProProValProProSerAlaThrGluProLeuAspLeuAspSerSe 184
QY 456 ACCTGTTGCC-----AACATTGGGCAACCCGAATTCCTCCCAATCTTA 500
Db 184 rSerCySerGlyThrProLeuHisArgGlnGlyProValGlnIleLeuProPheLeuTr 204
QY 501 TCTTGCTGCGACGAGATGCTCTCAACAGAGCTGATACAGACAGAAATGGATTGTTA 560
Db 204 rLeuGlySerAlaGlyHisAlaAlaArgArgAspMetLeuAspAlaLeuGlyIleThrAl 224
QY 561 TGTGTTAAAGCCAGCTATACCTGTCGAAGCTGACTTATCCCGAGTCAATTTCT 620
Db 224 aLeuLeuAsnValSerSerAspCyProAsn---HisPheGlyGlyHisTrpGlnTrpGly 243
QY 621 GCGTGGCTGTGTAATGACAGCTTTTGTGAAATAATTTTGCCTGTTGGAACAATCAGT 680
Db 243 sCyIleIleProValGlnAspAsnHisValAspIleSerSerTrpPheMetGlnAlaI 263
QY 681 AGATTTCATTGAGAAAGCAAGCTTCGAATGATGTTTCTAGTCACTGTTTAAGCTGG 740
Db 263 eGlnTrpIleLeuAspAlaValValAspCyArgGlyArgValIleValHisCysGlnAlaG 283

RESULT 10
US-09-671-325-805
Sequence 805, Application US/09671325
Patent No. 6667154
GENERAL INFORMATION:
APPLICANT: Wang, Tongfong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedrick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C12
CURRENT APPLICATION NUMBER: US/09/671,325
CURRENT FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 1825
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 805
LENGTH: 394
TYPE: PRT
ORGANISM: Homo sapiens
US-09-671-325-805

Alignment Scores:
Pred. No.: 1,25e-32 Length: 394
Score: 455.00 Matches: 122
Percent Similarity: 49.87% Conservative: 75
Best Local Similarity: 30.89% Mismatches: 143
Query Match: 12.59% Indels: 55
DB: Gaps: 13

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-671-325-805 (1-394)

QY 45 GGTGGCTGCTGCGAAAGTGAACGAA-----AAAGT 77
Db 26 GlyGlyAlaGly-SerGlySerHisGlyThrLeuGlyLeuProSerGlyGlySerCy 45
QY 78 GGTGTAATGTAAGCCGGCATTGTGGAATACATACATCCACATTTTGAAGCCAT 137
Db 45 GLeuLeuSerCyArgProPheLeuAlaHisSerAlaGlyTrpLeuGlySerVa 65
```

```

Qy 138 TAATATCACTGCTCCAGATGTTATGAGGAAAGTTGCAAGACGAAAGTTAAATTAC 197
Db 65 IAsnValArgCysAsnThrIleValArgArg---AlaYsgIlySerValSerIleuGI 84
Qy 198 AGAGCTC-----CAGCATTCAGCGAAACATTAAGTTGACATTGACAGTCAAGAA 251
Db 84 uGnIleLeuProAlaGIuGIuGIuValArgAlaArgLeuArgSerGlyLeuYrSerAl 104
Qy 252 GGTGTAGTTTACGATCAAGATCCCAAGAGTTGCTCTCTCTCTTCAAGACTGTTTCT 311
Db 104 aValIleValIYrAspGIuArgSerProAlaGIuSerIleuArgIuAspSerThrVa 124
Qy 312 CACTGTACTCTGGGTAACTGGAGAAAGCTTC-----AACTGTCTACCTGCTTGC 365
Db 124 IserIleuValIAlaIuAlaLeuArgArgAsnAlaGIuArgThrAspIleCysLeuLeu 144
Qy 366 AGTGGGTTTGTAGTCTCTGTTGTTTCCCTGGCCCTCTGGAAGAAATCCACTCT 425
Db 144 sGIYGIYrGIuArgPheSerSerGIuYrProGIuPheCysSerIYrThrIYsAlaIe 164
Qy 426 A-----GTCCCTACCTGCAATTCTCAGCT-----TGCTT 455
Db 164 uAlaAlaIleProProProProValProProSerAlaThrGIuProLeuAspLeuAspCysSe 184
Qy 456 ACCTGTGCC-----AACATGGGCGAACCGCAATTCTTCCCAATCTTTA 500
Db 184 rSerCysGIYrThrProLeuHIsAspGIuGIuGIuValArgIuLeuProPheLeuY 204
Qy 501 TCTTGGCTGCGACGAGATGCTCTCAACAGAGCTGATACAGAGATGTTGTTA 560
Db 204 rIeuGIYserAlaYrHIsAlaAlaArgArgAspMetLeuAspAlaLeuGIYleHrAl 224
Qy 561 TGTGTTAAATGCGCACTTACCTGTCCAAAGCTGACTTATCCCGAGTCTATTCT 620
Db 224 aLeuLeuAsnValIserSerAspCysProAsn---HisPheGIuGIYHIsIYrGIuYrIly 243
Qy 621 GGTGTGCTGTGATGACAGCTTTGTGAGAAATTTTGCCTGTGGAAACATTCAGT 680
Db 243 sCysIleProValGIuAspAsnHIsYsAlaAspIleSerSerThrPheMetGIuAlaI 263
Qy 681 AGATTTCAATTAGAAACAAAGCTTCATGAGATGTTGTTCAGTCTTAAAGCTGG 740
Db 263 eGIuYrIleAspAlaValIYsAspCysArgGIuValIleuValHIsCysGIuAlaGI 283
Qy 741 GATCTCCGCTCGCGCAACATCGCTATCGCTTACATCATGAGAGATGACATGTCTT 800
Db 283 YIleSerArgSerAlaThrIleCysLeuAlaYrIleuMetCetylSylsArgValArgLe 303
Qy 801 AGATGAAGCTTACAGATTGTGTGAAGAAAGAAAGCTTCTATATCTCCAAATTCAT 860
Db 303 uGIuGIuAlaPheGIuPheValIYsGIuArgIserIleIleSerProAsnPheSerPh 323
Qy 861 TCTGGGCCAACTCTGCTGATGAGAAAGATTTAAGAACCAAGCTGAGCATCAGGGCC 920
Db 323 eMetGIYGIuIleuLeuGIuPheGIuSerGIuValIleuAlaThrSerCysAlaIaGIuAl 343
Qy 921 AAAGGCAAACTCAAGCTGTGCACTTGAGAAAGCTCAAAATGAACTCTCCCTCTCTC 980
Db 343 aAlaSer-----ProSerGIYProLeu----- 350
Qy 981 AGAGGTGACAGAAAGCGAGAGCGCCCTCACTGCAACCTGTGGCGACTGTCTACCTC 1040
Db 351 -----GIYGIuArgGIYsIYrThrProAlaThrPro-----ThsSe 362
Qy 1041 AGAGGACGACGAGCAAAAGCCCGCTGATCCCGCAGAGGTGCGCAGGTGCCAGGTGCA 1100
Db 362 rGIuPheValPheSerPheProVal---SerValGIYAlaHIsSerAlaProSerSerIe 1101
Qy 1101 GCCGTGCTGTTTAGAGAGAGCCCGCTGCTGATCAGGCGCTCACT 1143
Db 381 uProYrIleu-----HisSerProIleThrThrIserProSer 393

```

```

US-09-589-184-805
/ Sequence 805, Application US/09589184
/ Patent No. 6686447
/ GENERAL INFORMATION:
/ APPLICANT: Mang, Tongtong
/ APPLICANT: Bangur, Chaltanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darriek
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ FILE REFERENCE: 210121.478C8
/ CURRENT FILING DATE: 2000-06-05
/ NUMBER OF SEQ ID NOS: 827
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 805
/ LENGTH: 394
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-589-184-805

Alignment Scores:
Pred. No.: 1,256-32 Length: 394
Score: 455.00 Matches: 122
Percent Similarity: 49.87% Conservative: 75
Best Local Similarity: 30.89% Mismatches: 143
Query Match: 12,598 Indels: 55
DB: 4 Gaps: 13

US-10-029-345a-108_copy_538_2532 (1-1995) x US-09-589-184-805 (1-394)
Qy 45 GGTGCTGCTGCGAGAAAGTTGGAAGCA-----AAAGT 77
Db 26 GIYGIYAlaGIYGIY-SerGIYSerHIsGIYrIleuGIYLeuProSerGIYGIYIYsCy 45
Qy 78 GCTGTAAATTATAGCGCGCATTTGTGGAATACATACATCCCATTTTGAAGCCAT 137
Db 45 eIleuLeuAspCysArgProPheLeuAlaHIsSerAlaGIYrIleuGIYSerVa 65
Qy 138 TAATATCACTGCTCCAGATGTTATGAGGAAAGTTGCAAGACGAAAGTTAAATTAC 197
Db 65 IAsnValArgCysAsnThrIleValArgArg---AlaYsgIlySerValSerIleuGI 84
Qy 198 AGAGCTC-----CAGCATTCAGCGAAACATTAAGTTGACATTGACAGTCAAGAA 251
Db 84 uGnIleLeuProAlaGIuGIuGIuValArgAlaArgLeuArgSerGlyLeuYrSerAl 104
Qy 252 GGTGTAGTTTACGATCAAGATCCCAAGAGTTGCTCTCTCTCTTCAAGACTGTTTCT 311
Db 104 aValIleValIYrAspGIuArgSerProAlaGIuSerIleuArgIuAspSerThrVa 124
Qy 312 CACTGTACTCTGGGTAACTGGAGAAAGCTTC-----AACTGTCTACCTGCTTGC 365
Db 124 IserIleuValIAlaIuAlaLeuArgArgAsnAlaGIuArgThrAspIleCysLeuLeu 144
Qy 366 AGTGGGTTTGTAGTCTCTGTTGTTTCCCTGGCCCTCTGGAAGAAATCCACTCT 425
Db 144 sGIYGIYrGIuArgPheSerSerGIuYrProGIuPheCysSerIYrThrIYsAlaIe 164
Qy 426 A-----GTCCCTACCTGCAATTCTCAGCT-----TGCTT 455
Db 164 uAlaAlaIleProProProProValProProSerAlaThrGIuProLeuAspLeuAspCysSe 184
Qy 456 ACCTGTGCC-----AACATGGGCGAACCGCAATTCTTCCCAATCTTTA 500
Db 184 rSerCysGIYrThrProLeuHIsAspGIuGIuGIuValArgIuLeuProPheLeuY 204
Qy 501 TCTTGGCTGCGACGAGATGCTCTCAACAGAGCTGATACAGCAATGGATGGATTGTTA 560

```

```

Dh 204 rLeuGIySerAlaTYrHIsAlAlAArgArGArSerMetLeuSerAlaLeuGIyIlleThrAl 224
Qy 561 TGtGTTAAATGCCAGCTACTACTGTGTCCAAAGCCCTACTTATATCCCGAGTCTCATTTCT 620
Dh 224 aLeuLeuAnValSerSerAspCyArProAsn---HisPheGluGIyHIsTYrGIyTrIly 243
Qy 621 GCGTGTGCTGTGAATGACAGCTTTTGTGTGAAGAAATTTTGGCCGTGGTGGACAAATTCAGT 680
Dh 243 eCyAlleProValGIuApeSerHIsLYeAlAlAspIleSerSerTrpPheMetGIuAlAl 263
Qy 681 AGATTTCATTTGAGAAAGCAAAAGCCCTCCAAATGATGTGTTCCTAGTCACATGTTTAgCTGG 740
Dh 263 eGluTYrIleApeAlaValLYeAspCyArGIyArGIyAlValLeuValHIsCyGIuAlAl 283
Qy 741 GATCTCCCGCTCCGCCACCATGTGCTATGCGCTTACATCATGAAGAGATGACATGCTTT 800
Dh 283 yIleSerArGIserAlaThrIlleCyAleuAlaTYrIleuMetMetLYsArGIyAlArGIe 303
Qy 801 AGATGAACCTTACACATTTTGTGAAGAAAGAAAGACCTTACTATATCTCCAAATTCATTT 860
Dh 303 uGluGIuAlaPheGIuPheValLYeGIuArGIyArGIserIlleIleSerProAsnPheSerH 323
Qy 861 TCTGGGCGCAACTCTGTGACTATGATGAAGAAGATTTAAAGAACACGACTGAGCATCGAGCC 920
Dh 323 eMetCIyInleuLeuGIuInPheGIuSerGIuAlValleuAlaThrSerCyAlAlAlGIuAl 343
Qy 921 AAAGAGCAAACTCAAGCTGTGCACCTGTGAGAGAACCCAAATGAACCTGTCCGTGTCTC 980
Dh 343 aAlaSer-----ProSerGIyPheLeu----- 350
Qy 981 AGAGGTGTGACAGAAAGCGAGACGCCCTCTACGTCCACCTGTGCCACTGTGTACTCTC 104
Dh 351 -----GIyGIuArGIyGIyThrProAlaThrPro-----ThrSe 362
Qy 1041 AGAGCAGCAGAGCAAAAGCCCGCTGCATCCCGCCAGCGTGGCCAGAGTGGCCAGCGTGA 1100
Dh 362 rGIuInPheAlaPheSerHIsProVal---SerValGIyAlaHIsSerAlaIleProSerSerIe 381
Qy 1101 GCCGTGCTGTTAAGAGACAGCCCGCGCTGTGTACAGCGCTCACT 1143
Dh 381 uProTYrIleu-----HisSerProIleThrThrSerProSer 393

RESULT 12
US-08-530-290-23
: Sequence 23, Application US/08530290
: Patent No. 5958721
: GENERAL INFORMATION:
: APPLICANT: Marshall, Christopher John
: APPLICANT: Ashworth, Alan
: APPLICANT: Hughes, David Anthony
: TITLE OF INVENTION: Methods for Screening of Substances for
: TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/530,290
: FILING DATE: 14-DEC-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/GB94/00694
: FILING DATE: 31-MAR-1994
: PRIOR APPLICATION DATA:

```

```

1  APPLICATION NUMBER: GB 9402573.1
2  FILING DATE: 10-FEB-1994
3  PRIOR APPLICATION DATA:
4  APPLICATION NUMBER: 9307250.2
5  FILING DATE: 07-APR-1993
6  ATTORNEY/AGENT INFORMATION:
7  NAME: Baetian, Kevin L.
8  REGISTRATION NUMBER: 34,774
9  REFERENCE/DOCKET NUMBER: 086611-0000000US
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE: (415) 576-0200
12 TELEFAX: (415) 576-0300
13 INFORMATION FOR SEQ ID NO: 23:
14 SEQUENCE CHARACTERISTICS:
15     LENGTH: 394 amino acids
16     TYPE: amino acid
17     STRANDEDNESS:
18     TOPOLOGY: linear
19     MOLECULE TYPE: peptide
20 US-08-530-290-23
21
22 Alignment Scores:
23 Pred. No.: 6,62e-32      Length: 394
24 Score: 447.00           Matches: 121
25 Percent Similarity: 49.62%      Conservative: 75
26 Best Local Similarity: 30.63%    Mismatches: 144
27 Query Match: 12.37%             Indels: 55
28 DB: 2                       Gaps: 13
29
30 US-10-029-345A-108_COPY_538_2532 (1-1995) x US-08-530-290-23 (1-394)
31
32 QY 45 GGTGGCTCTGCTGGAAAGTGAACGGA-----AAAGT 77
33 Db 26 GYGLYGLIAGLYGLY-SerGIYserThIeGLYThIeGLYLeuProSerGIYGLYLeuCY 45
34
35 QY 78 GCGCAATAGTAATAGCGCGGCATTTGTGGAAATACATACATCCACATTTTGGAAAGCAT 137
36 Db 45 sIeuLeuLeuAerCYsArGrProPhLeuAlaHiserIaGLYTrIleuGLYserVa 65
37
38 QY 138 TAATATCAACTGCTCCCAAGCTTATGAAAGCGAGGTTCGACAGACAAAGTGTAAATTAC 197
39 Db 65 IAnVaLaVArGYsAenThIeValArGrArG---AlaLeuGIserValserLeuGI 84
40
41 QY 198 AGAGCTCATC-----CAGCATTCAGCGAAACATPAGGTTGACATTGATGGCACTCGAA 251
42 Db 84 uGIInIleuProAlaGLIuGLIuValArAlaArGrLeuArGrserGIYLeuTrYserAl 104
43
44 QY 252 GGTGTAGTTTACGATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCAGACTGTTTCT 311
45 Db 104 aValIleValIYArGrGIuArGrserProArGAlaGLIserLeuArGrGIuArGrserThrVa 124
46
47 QY 312 CAGTACTACTTGGGGTAATGAAAGAGCTTC-----AACTGTGTCACCTGTTCG 365
48 Db 124 IserLeuValIValGIAlaLeuArGrArGrAlaGLIuArGrThArGrIleuArGrIleuLeu 144
49
50 QY 366 AGGTGGGTTGTGAGGTTCTCTGTGTGTTTCCCTGGCCCTGTGAAGAAATCCACTCT 425
51 Db 144 sGIYGLYTrYrGIuArGrPheserSerGIuTrProGIuPheserIYserThrIYsAlaLe 164
52
53 QY 426 A-----GTCCCTAACCTGCATTTCTCAGCT-----TCCTT 455
54 Db 164 uAlaAlaIleProProProValProProSerAlaThrGIuProLeuArGrLeuArGrCyse 184
55
56 QY 456 ACCTGTGGC-----ACATTTGGGCAACCGGAATTTCTTCCCAATCTTTA 500
57 Db 184 rSerCYsGIYThrProLeuHisArGrInGIuGIYProValGIuIleuLeuProPherLeu 204
58
59 QY 501 TCTTGGCTGCGCAGCAGATGTCCTTCACAAAGAGCTGATACAGCAGATGGATGGTTA 560
60 Db 204 rIleuGIYserAlaTrYrHisAlaAlaArGrArGrArGrArGrArGrArGrArGrArGr 224
61
62 QY 561 TGCTGTTAAATGCCAGCTATACCTGTGCCAAAGCTGACCTTTATCCCGAGTCTCATTTCT 620
63 Db 561

```

```

Db      224 aLeuLeuAenValSerSerAspCysProAsn-----HisPheGluGlyHisTyrglnTyrIly 243
QY      621 GCCTGTGCGCTGGTGAATGACGCTTTTGTGAGAAAATTTTGGCGGTGTGACAAATCAGT 680
Db      243 sCysIleProValIGluAspAsnHisLySAlaAspIleSerSerTrpPheMetGluAlaIi 263
QY      681 AGATTTCATTGAGAAAGCAAAAGCCCTCAATGATGATGTGTTCTAGTGCATCTTTAGCTGG 740
Db      263 eGluTyrlleAspAlaValIlyAspCysArgGlyArgValIleuValHisCysGlnAlaGl 283
QY      741 GATCTCCGCTCCGCGCACATCGCTATCGCTTACATCATGAGAGATGAGATGATGCTTT 800
Db      283 yIleSerArgSerAlaThrIleCysleuAlaTyrlleuMetLysLyArgValArgIe 303
QY      801 AGATGAAGCTTACAGATTGTGAAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT 860
Db      303 uGlnGluAlaPheGluInPheGluInPheGluInValIleuAlaThrSerCysAlaIaGluAl 323
QY      861 TCTGGGCGCAACTCCTGGACATATGAGAAAGATTAAAGAACAGACTGGAGATCAGAGGCC 920
Db      323 eMetGlyGlnleuLeuGlnPheGluInPheGluInValIleuAlaThrSerCysAlaIaGluAl 343
QY      921 AAAGAGCAAACTCAAGCTGCTGCACTGTGAGAAAGCAATGAACTGCTGCTGCTGCTC 980
Db      343 aAlaSer-----ProSerGlyProLeu----- 350
QY      981 AGAGGCTGACAGAAAAGGAGAGAGCCCTCAGTCCACCTGTCGCCAGCTGTGCTACCTG 1040
Db      351 -----ArgGluArgGlyLySAlaThrProAlaThrPro-----Thrse 362
QY      1041 AGAGGAGAGAGAGCAAAAGCCGCTGATCCCGCAGCGTGGCCAGCGTGGCCAGCGTGA 1100
Db      362 tGlnPheValIlePheSerPheProVal---SerValGlyValHisSerAlaProSerSerle 381
QY      1101 GCCGTGCTGTGTTAGAGACAGCCGCTGTGATAGAGCCCTCAGT 1143
Db      381 uProTyrlleu-----HisSerProIleThrTrpSerProSer 393

RESULT 13
US-09-702-705-827
; Sequence 827 Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedrick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: DIAGNOSIS OF LUNG CANCER
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 827
; LENGTH: 394
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-702-705-827

```

Alignment Scores:

```

Pred. No.: 6.62e-32 Length: 394
Score: 447.00 Matches: 122
Percent Similarity: 50.38% Conservative: 77
Best Local Similarity: 30.89% Mismatches: 141
Query Match: 12.37% Indels: 55
DB: 4 Gaps: 14

```

```

US-10-029-345a-108_copy_538_2532 (1-1995) x US-09-702-705-827 (1-394)
QY      45 GGTGGCTGCTGGTGAAGAGGAGCAAGAA-----AAAGT 77
Db      26 GIGlyAlaGlyGly-SerGlySerHisSglYThrLeuGlyLeuProSerGlyGlyIyScy 45
QY      78 GCTGCTAAATTGATTAACCGCGCATTTTGTGATGATCAATACATCCACATTTTGGAAACCAT 137
Db      45 sIleuLeuLeuAspCysArgProPheLeuAlaHisSerAlaGlyTyrlleuGlySerVa 65
QY      138 TAATACACAGCTGCTCCAGCTTATGAGACCGAAGGTTCACACAGACAAAGTGTATATAC 197
Db      65 lAenValArgCysAsnThrIleValAlaArgArgArg--AlaIySglYSerValSerLeuGl 84
QY      199 AGAGCTATC-----CAGCATTCAGCGAAACATAAGTTGACATTGATTCAGACAGAA 251
Db      84 uGlnIleuProAlaGlnIleuGluValAlaArgAlaArgLeuArgSerGlyLeuTyrlSerAl 104
QY      252 GGTGTAGATTATGATGAAAGCTCCCAAGATGTTGCTCTCTCTTCCAGACTGTTTCT 311
Db      104 aValIleValIlyThrPheIlyArgSerProAlaArgAlaGlnIleuArgHisPheSerThrVa 124
QY      312 CACTGTCTCTTGGGTAAACTGAGAGAGAGCTTC-----AACTGTTCACCTGCTTGC 365
Db      124 lSerIleuValValGlnAlaIleuArgArgAsnAlaGlnArgHisPheIlyCysleuLeuLy 144
QY      366 AGGTGGATTGTGAGATTCTCTGTTTCCCTGCGCTGTGAGAGAAATCCACTCT 425
Db      144 sGlyGlyTyrgluAlaPheSerSerSerGlyArgProGlnPheCysSerIyThrIySalale 164
QY      426 A-----GTCCCTACCTGCATTTCTCAGCCT----- 450
Db      164 uAlaAlaIleProProProValProProSerAlaThrGlnProLeuAspIleuGlyCysSe 184
QY      451 ----TGC--TTACCTGTGGCCAACTT---GGGCCAACCCGAATTTCTCCAACTTTA 500
Db      184 rSerCysGlyThrProLeuHisAspGlnGlyGlyProValIGluIleuProPheLeuTy 204
QY      501 TCTTGGCTGCCAGAGATGTCTTCAACAGAGAGATGATACAGCAGAAATGAGATTGGTTA 560
Db      204 tIleuGlySerAlaTyrlHisAlaAlaArgArgAspMetLeuAspAlaIleuGlyIleThrAl 224
QY      561 TGTGTTAATGCCAGATATACCTGTGCCAAAGCTGATTTATCCCGAGTCAATTTCT 620
Db      224 aLeuLeuAsnValSerSerAspCysProAsn---HisPheGluGlyHisTyrglnTyrIly 243
QY      621 GCGTGTGCTGTGATGACAGCTTTTGTGAGAAAATTTTGGCGGTGTGACAAATCAGT 680
Db      243 sCysIleProValIGluAspAsnHisLySAlaAspIleSerSerTrpPheMetGluAlaIi 263
QY      681 AGATTTCATTGAGAAAGCAAAAGCCCTCAATGATGATGTGTTCTAGTGCATCTTTAGCTGG 740
Db      263 eGluTyrlleAspAlaValIlyAspCysArgGlyArgValIleuValHisCysGlnAlaGl 283
QY      741 GATCTCCGCTCCGCGCACATCGCTATCGCTTACATCATGAGAGATGAGATGATGCTTT 800
Db      283 yIleSerArgSerAlaThrIleCysleuAlaTyrlleuMetLysLyArgValArgIe 303
QY      801 AGATGAAGCTTACAGATTGTGAAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGT 860
Db      303 uGlnGluAlaPheGluInPheGluInPheGluInValIleuAlaThrSerCysAlaIaGluAl 323
QY      861 TCTGGGCGCAACTCCTGGACATATGAGAAAGATTAAAGAACAGACTGGAGATCAGAGGCC 920
Db      323 eMetGlyGlnleuLeuGlnPheGluInPheGluInValIleuAlaThrSerCysAlaIaGluAl 343
QY      921 AAAGAGCAAACTCAAGCTGCTGCACTGTGAGAAAGCAATGAACTGCTGCTGCTGCTC 980
Db      343 aAlaSer-----ProSerGlyProLeu----- 350
QY      981 AGAGGCTGACAGAAAAGGAGAGAGCCCTCAGTCCACCTGTCGCCAGCTGTGCTACCTG 1040
Db      351 -----ArgGluArgGlyLySAlaThrProAlaThrPro-----Thrse 362

```



```

OY      1041 AGAGGACACAGAGACAAAGCCCGTCATCTCCGACGCGTCCCAAGCTGCCCAAGGTGCA 1100
      ||:::|||||
DB      362 TGLnheValPheSerPheProter---SerValcIyValHisSerAlaProSerSerle 381
OY      1101 GCGCGCGCTGTTAAGAGACAGCCCGCTGTACAGCGCTCAT 1143
      ||| ||| |||||::: |||
DB      381 uProtyrLeu-----HisSerProtlethrTherProser 393

RESULT 14
US-09-736-457-827
; Sequence 827, Application US/09736457
; Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaifanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darlick
APPLICANT: Retter, Marc
APPLICANT: Manlon, Jane
APPLICANT: Fan, Jiqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT FILING DATE: US/09/736.457
NUMBER OF SEQ ID NOS: 2000-12-13
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 827
LENGTH: 394
TYPE: PRT
ORGANISM: Homo sapiens
US-09-736-457-827

Alignment Scores:
Pred. No.: 6.62e-32 Length: 394
Score: 447.00 Matches: 122
Percent Similarity: 50.38% Conservative: 77
Best Local Similarity: 30.89% Mismatches: 141
Query Match: 12.37% Indels: 55
Gaps: 4
DB: 14

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-736-457-827 (1-394)
OY      45 GGTGGCTCTGTGGAAAGTGAACGGA-----AAAGT 77
      |||||:::|||||
DB      26 GLyGlyAlaGlyGly-SerGlySerHisGlyThrLeuGlyLeuProSerGlyGlyLeuCy 45
OY      78 GCTGCTATTGATAGCCGCGCATTTGTGTGAATACATACATCCCACTTTTGGAAAGCAT 137
      |||||:::|||||
DB      45 sLeuLeuLeuAspCyAspArgProPheLeuAlaHisSerAlaGlyUtyrLeuGlySerVa 65
OY      138 TAATATCAACGCTCCCAAGCTTATGACGGAAGGTTCAGACAGCAAGAAAGTTATATAC 197
      |||||:::|||||
DB      65 lAsnValArgCyAsnThrIleValArgArgG---AlaValGlySerValSerLeuG1 84
OY      198 AGAGCTCATC-----CAGCATTCAGCAAACTAAGTTGACATTTGCACTGACAGAA 251
      |||||::: |||
DB      84 uGlnIleLeuProAlaGluGluGluValArgAlaArgLeuAspArgSerGlyLeuTySerAl 104
OY      252 GGTGTGATTATACATCAAGCTCCCAAGATGTTCCCTCTCTTTCAGACTGTTTCT 311
      |||||:::|||||
DB      104 aValIleValIlyrAspGluArgSerProArgAlaGluSerLeuAlaGluAspSerThVa 124
OY      312 CACTGTACTTGTGGGTAACTGAGAAAGACTTC-----AAGTCTGTTCACCTGCTTGC 365
      |||||::: |||
DB      124 lSerLeuValValGlnAlaIleuArgAspAlaGluArgThrAspIleCyValLeuLeuLy 144
OY      366 AGTGGGTTTGCTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGAAATTCATCT 425
      |||||::: |||||

```

```

Db      144  sGlyGlyTYGlyGlnArgHseSerSerGluTYrProGlnPheCysSerIyThrIySAlaLe 164
Oy      426  A-----GTCCCTACTGTCATTTCTCAGCCT-----450
Db      164  uAlaAlaIleProProProValProProSerAlaThrGlnProLeuAlaPheGlyCysE 184
Oy      451  -----TGC---TTACTGTGGCGCAACATT---GGGCCAACCCGAATTTCTCCGAATCTTTA 500
Db      184  rSerCysGlyThrProLeuHisArgGlnGlyGlyProValGlnIleuProPheLeuTY 204
Oy      501  TCTTGCGTCGCAGCAGATGTCTTCACCAAGAGCTGATACAGCAGATGGATTTGGTTA 560
Db      204  rLeuGlySerAlaTyThrIleAlaAlaAlaArgAlaGArgPheMetLeuAlaLeuGlyIleThrAl 224
Oy      561  TGTTTAAATGCGACGCTATACCTGTCCAAAGCCTGACTTATATCCCGAGTCTCATTTCCCT 620
Db      224  AlaLeuAlaHisValSerSerArgCysProAla---HisPheGlnGlyHisIleTYrGlnTYrLY 243
Oy      621  GCGTGTCGCTGTGAATGACAGCTTTTGTGTGAAGAAATTTTGGCGGTGGTGGCAAAATGAGT 680
Db      243  sCysIleProValGlnAlaPheHisIySAlaAlaPheIleSerSerIlePheMetCysAlaAlaI 263
Oy      681  AGATTTGATTTGAGAAACAAAGCCCTCAATGGATGTGTTTGTAGTCAGCTTTTACCTGG 740
Db      263  eGluTYrIleAspAlaValIySAspCysArgGlyArgValIleValHisCysGlnAlaG 283
Oy      741  GATCTCCGCTCGGCCACCATTCGCTATGCGCTTACATCATGAAGAAGATGACATGTCTTT 800
Db      283  yIleSerArgSerAlaThrIleCysAlaAlaTYrLeuMetMetIySArgValAlaTArgLe 303
Oy      801  AGATGAAGCTTACGATTTTGTGAAAGAAAAAAGACCTACATATATCCCAACCTCATTT 860
Db      303  uGlnGlnAlaPheGlnPheGlnValIySAlaArgSerIleIleSerProAlaPheSerPh 323
Oy      861  TCTGGGCAACTCCTGTGACTATGAGAAGATTAAGAACACAGACTGAGCATCAGAGCC 920
Db      323  eMeGlyGlnLeuGlnPheGlnSerGlnValIleAlaThrSerCysAlaAlaGlnAl 343
Oy      921  AAAGAGCAAACTCAAGCTGTGCACCTGAGAAAGCCAAATGAATGATCTCCCTGTGTCTC 980
Db      343  aAlaSer-----ProSerGlyProLeu-----350
Oy      981  AGAGGGTGACACAAAAAGCCAGAGGCCCTTCAGTCCACCTGTGCGCATCTGTACTCTC 1040
Db      351  -----ArgGlnArgGlyIySThrProAlaThrPro-----ThrSe 362
Oy      1041  AGAGGACAGACAGCAAAAGGCCCGGATCCCGCAGCGTCCGCGCAGCGTGCAGTCA 1100
Db      362  rGlnPheValIlePheSerPheProVal---SerValGlyValHisSerAlaProSerSerLe 381
Oy      1101  GCCGTGCTGTGTAGAGGACAGCCCGCTGTATACAGCGCCTCAGT 1143
Db      381  uProTYrLeu-----HisSerProIleThrIleThrSerProSer 393

RESULT 15
US-09-614-124B-827
: Sequence 827, Application US/09614124B
: Patent No. 6630574
: GENERAL INFORMATION:
: APPLICANT: Wang, Tongtong
: APPLICANT: Bangur, Chaitanya S.
: APPLICANT: Lodes, Michael A.
: APPLICANT: Fanger, Gary
: APPLICANT: Veddvick, Tom
: APPLICANT: Carter, Darriick
: APPLICANT: Retter, Marc
: APPLICANT: Mannion, Jane
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
: TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
: FILE REFERENCE: 210121.478C9
: CURRENT APPLICATION NUMBER: US/09/614,124B
: CURRENT FILING DATE: 2001-07-11
: NUMBER OF SEQ ID NOS: 1668

```

```

; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 827
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B-827

Alignment Scores:
Pred. No.: 6,62e-32      Length: 394
Score: 447.00           Matches: 122
Percent Similarity: 50.38%      Conservative: 77
Best Local Similarity: 30.89%    Mismatches: 141
Query Match: 12.37%           Indels: 55
DB: 4                      Gaps: 14

US-10-029-345A-108_copy_538_2532 (1-1995) x US-09-614-124B-827 (1-394)

QY 45 GGTGGCTGCTGGAAGAGGAAACGGA-----AAAGT 77
DB 26 GlyGlyAlaGlyGly-SerGlySerHisGlyThrLeuGlyLeuProSerGlyGlyLeuCy 45
QY 78 GGTGCTAATTGATAGCCGGCATTTGTGGATACATACATCCACATTTTGGAGCCAT 137
DB 45 sLeuLeuLeuAspCysArgProPheLeuLHisSerAlaGlyTyrLeuGlySerVa 65
QY 138 TAAATCAACTGCTCCAGCTTATGAGCGAAGGTTCACACGAAAGTGTAAATAC 197
DB 65 lAsnValArgCysAsnThrIleValArgArgArg--AlaLysGlySerValSerLeuG 84
QY 198 AGAGCTCATC-----CAGCATTCAGCGAAACATTAAGTTGACATTGATGCAGTCA 251
DB 84 uGlnIleLeuProAlaGluGluGluValArgAlaArgLeuArgSerGlyLeuTyrSerAl 104
QY 252 GGTGTAGTTAGTACATCAAGCTCCCAAGATGTGCTCTCTCTCTGACACTGTTTCT 311
DB 104 aValIleValTyrAspGluArgSerProArgAlaGluSerLeuArgGluAspSerThrVa 124
QY 312 CACTGACTCTTGGGTAAACTGGAGAAGACTT-----AACTGTGTTCACTGCTTGC 365
DB 124 lSerLeuValValGlnAlaLeuArgAsnAlaGluArgThrAspIleCysLeuLeuLy 144
QY 366 AGGTGGTTTGCTGAGTCTCTCGTGTTCCTGGCTGCTGCTGGAAGAAATCCACTCT 425
DB 144 sGlyGlyTyrGluArgPheSerSerGlyTyrProGluPheCysSerLysThrLysAlaLe 164
QY 426 A-----GTCCCTAAGCTGCAATTTCTCAGCT----- 450
DB 164 uAlaIleIleProProProValProProSerAlaThrGluProLeuAspLeuGlyCysSe 184
QY 451 -----TGC---TTACCTGTGGCCCAACT--GGCCAAACCGAATTTTCCCAATCTTTA 500
DB 184 lSerCysGlyThrProLeuHisAspGlnGlyGlyProValGluIleLeuProPheLeuTy 204
QY 501 TCTTGGCTGCAGGAGATGCTCAACAGGAGCTGATACAGAGCAATGGGATGGTTA 560
DB 204 rLeuGlySerAlaTyrHisAlaAlaArgArgAspMetLeuAspAlaLeuGlyIleThrAl 224
QY 561 TGTGTTAAATGCGACCTATACCTGTGCAAGCTGACTTATCCCGAGTCTCATTTTCT 620
DB 224 aLeuLeuAsnValSerSerAspCysProAsn--HisPheGluGlyHisTyrGlnTyrLy 243
QY 621 GCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTGGCCGTGGTTGACAAATCAGT 680
DB 243 sCysIleProValGluAspAsnHisLysAlaAspIleSerSerTrpPheMetGlnAlaI 263
QY 681 AGATTTCATTGAGAAAGAAAGCCTCAATGATGATGTCTTAGGCACTGTTAGCTGG 740
DB 263 eGlnTyrIleAspAlaValLysAspCysArgGlyArgValIleuValHisCysGlnAlaG 283
QY 741 GATCTCCGCTCCGCGACCATCGCTATGCGCTATCATGAGAGATGAGCATGCTCTT 800
DB 283 yLlSerArgSerAlaThrIleCysLeuAlaTyrLeuMetMetLysLysArgValArgLe 303

```

```

QY 801 AGATGAAGCTTACAGATTGTGAAAGAAAGAAAGCCTACTATATCTCCAAACTTCATTT 860
DB 303 uGlnGluAlaIlePheGluPheValLysGlnArgArgSerIleIleSerProAsnPheSerPh 323
QY 861 TCTGGGCCAACTCCTGGACTATGAGAAAGATTTAAGAACCAAGCTGAGCATCAGGGCC 920
DB 323 eMetGlyGlnLeuLeuGlnPheGluSerGlnValIleuAlaThrSerCysAlaIaGluAl 343
QY 921 AAAGAGCAACTCAGCTGCTGACCTCGAGAAAGCCAAATGAACCTGCTCCTGCTCTC 980
DB 343 aAlaSer-----ProSerGlyProLeu-----ThrSe 362
QY 981 AGAGGTGACAGAAAGCGAGACGCCCTCACTGTCACCTGTGCGGACTGTGCTACCTC 1040
DB 351 -----ArgGluArgGlyLysThrProAlaThrPro----- 362
QY 1041 AGAGGACAGCAGACAAAGCCCGTGCATCCCGCAGCGTCCCAAGCTGCCAGCTGCA 1100
DB 362 rGlnPheValPheSerPheProVal--SerValGlyValHisSerAlaProSerSerLe 381
QY 1101 GCGGTGCTGTTAGAGGACAGCGCGCTGTACAGGGCGCTCAGT 1143
DB 381 uProTyrLeu-----HisSerProIleThrThrSerProSer 393

```

Search completed: June 21, 2004, 12:43:22
 Job time : 42.4447 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 12:17:10 ; Search time 12558 Seconds

(without alignments)
6885.611 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccaatgagatgattg9.....tggaatcatgagctctcc 1995

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: *
1: gb_ba: *
2: gb_hg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_mu: *
19: em_om: *
20: em_ov: *
21: em_or: *
22: em_ov: *
23: em_ph: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hg_hum: *
31: em_hg_inv: *
32: em_hg_other: *
33: em_hg_mus: *
34: em_hg_pln: *
35: em_hg_rod: *
36: em_hg_mam: *
37: em_hg_vrt: *
38: em_sy: *
39: em_hggo_hum: *
40: em_hggo_mus: *
41: em_hggo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1995	100.0	5450	6	AX482439
2	1995	100.0	5450	6	AX482478
3	1991.8	99.8	1998	6	AX260342
4	1991.8	99.8	3059	6	AX278461
5	1991.8	99.8	3496	6	AX441210
6	1991.8	99.8	3521	9	AB052156
7	1991.8	99.8	3544	6	AX260340
8	1991.8	99.8	3566	9	AF506796
9	1991.8	99.8	3766	6	AX374994
10	1991.8	99.8	4790	6	BD183422
11	1991.8	99.8	4790	6	AB051487
12	1991.8	99.8	4790	9	AB051487
13	1990.2	99.8	2732	6	AX180875
14	1990.2	99.8	3104	6	AX405700
15	1988.6	99.7	2102	6	AX713989
16	1988.6	99.7	2102	9	AK055973
17	1988.6	99.7	5111	6	AX482372
18	1988.6	99.7	2118	6	AX099933
19	1963.8	98.4	3284	6	BC042101
20	1962.2	98.4	2071	6	AX921917
21	1906.8	95.6	2200	6	AX921915
22	1844.2	92.4	2807	9	BC031643
23	1653.8	82.9	1935	9	AY038927
24	1653.8	82.9	3332	6	AX441229
25	1572.4	78.8	4943	10	BC059232
26	1572.4	78.8	4975	10	BC057321
27	1567.6	78.6	4874	10	AB052157
28	1554.8	77.9	2756	6	AX482444
29	1531	76.7	4992	10	AF345951
30	1200.6	60.2	4827	10	AF345952
31	1183.4	59.3	172206	9	AC007619
32	1183.4	59.3	188344	2	AC131617
33	879.4	44.1	1916	6	AX835305
34	879.4	44.1	1916	6	AX098310
35	858.4	43.0	201474	2	AC126692
36	858.4	43.0	236589	2	AC118035
37	848.6	42.5	242590	2	AC133722
38	848.6	42.5	244605	2	AC097818
39	848.6	42.5	250782	2	AC128093
40	848.6	42.5	250782	6	AX868469
41	727.6	36.5	749	6	BD148531
42	714.2	35.8	4026	10	AF345953
43	405.8	20.3	3231	5	BC063941
44	387.8	19.4	2377	6	AR255944
45	387.8	19.4	2377	9	HS027193

ALIGNMENTS

RESULT 1
AX482439
LOCUS AX482439 5450 bp DNA
DEFINITION Sequence 108 from Patent WO02057460.
ACCESSION AX482439
VERSION AX482439.1 GI:22316984
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Todderud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G.,
Suchard,S., Banas,D., Baesolino,D., Feder,J., Krystek,S.,
Meatee,P., Mintier,G., Siemers,N., Jackson,D.G. and Ramanathan,C.

Pred. No. is the number of results predicted by chance to have a

TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 108 25-Jul-2002;
SQUIBB BRISTOL MYERS CO (US)
FEATURES
source location/Qualifiers

1..5450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538..2535
/note="unamed protein product"

/codon_start=1
/protein_id="CAD4451.1"
/db_xref="GI:22316385"

/translation="NAHEMIGTOIVTERVILLESSETEKVLIDSRPFVEYNTSHILE
AIVNCSKMKRRLQODKVLITBLIDHSKHKVDIDCSQKVVVDSDSDVASLSDDC
FLVLVGLKLEKSNVHLIAGGPAEBSRFPGLCEGKSTLTPFCISQCLPVAIVGTT
RLPNIYLQGDQVLDVNLKELIQNGIGVYLMASYCPKEDPIFESHPLRPVNDSPCEK
KRPITSPNPNFLGQULIDYEKIKNOTGASGPKSKLILHEKPNRPVPAVSEGSQSE
TNLSPKADSDATSEAAAGRPVHPASVSPVOPSLDESLVVALGSLHLSADRLD
LSQPEPSPDKERASIPKLOTPAPSPDSQSKLHVSRTSSGTAGRNKLCQSPVOE
VEDNHTSFLPGLSTSOHLTKSAGLGMHSDILAPOTSBLTISMTYFATRESHP
YSASAIYGGASYSYASCSQPLTCGDQVYSVRRQKSDSDADSRSMWHESSPEKQFK
RRSCMEBESITSENRSREBELKRVGSQSPSGSMETIEVS"

ORIGIN

Query Match 100.0%; Score 1995; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGCCCATGAGATGATTTGGAACCTCAATTGTTACTGAGAGTTGGTCTCTGTGAA 60
DB 538 ATGCCCATGAGATGATTTGGAACCTCAATTGTTACTGAGAGTTGGTCTCTGTGAA 60
QY 61 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCGCGCATTTGTGAATCAATATCATCC 120
DB 598 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCGCGCATTTGTGAATCAATATCATCC 120
QY 121 CACATTTTGAAGCATTAATATCACTGCTCCAACTTATGAGGGAAGTTGCAACG 180
DB 658 CACATTTTGAAGCATTAATATCACTGCTCCAACTTATGAGGGAAGTTGCAACG 180
QY 181 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
DB 718 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
QY 241 TGCAGTCAGAAAGGTGTAGTTTACGATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB 778 TGCAGTCAGAAAGGTGTAGTTTACGATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCA 300
QY 301 GACTGTTTTCTCACTGTACTCTTGAGGTAACCTGAGAGAGAGCTTCAACTCTGTTCACTG 360
DB 838 GACTGTTTTCTCACTGTACTCTTGAGGTAACCTGAGAGAGAGCTTCAACTCTGTTCACTG 360
QY 361 CTTCGAGGAGGTTTGTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGGAAATCC 420
DB 898 CTTCGAGGAGGTTTGTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGGAAATCC 420
QY 421 ACTCTAGTCCCTACCTGCAATTCAGAGCTTGTACTCTGTTGCAACATTTGGGCCAAC 480
DB 958 ACTCTAGTCCCTACCTGCAATTCAGAGCTTGTACTCTGTTGCAACATTTGGGCCAAC 480
QY 481 CGAATTTCTCCCATCTTTATCTTGGCTGCAAGAGATGCTCTCAACAAGAGAGCTGATA 540
DB 1018 CGAATTTCTCCCATCTTTATCTTGGCTGCAAGAGATGCTCTCAACAAGAGAGCTGATA 540
QY 541 CAGAGAAATGGGATTTGTTATGTTTAAATGCAAGCATTAATCTGTGCAAAAGCTTGAAT 600
DB 1078 CAGAGAAATGGGATTTGTTATGTTTAAATGCAAGCATTAATCTGTGCAAAAGCTTGAAT 600
QY 601 ATCCCGAGTCTAATTTCTGCGGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 660
```

```
DB 1138 ATCCCGAGTCTAATTTCTGCGGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 1197
QY 661 CCGAGTTTGGACAAATCAGTAAATTTCAATTGAAAGCAAAAGCTCCATGATGTTGTT 720
DB 1198 CCGAGTTTGGACAAATCAGTAAATTTCAATTGAAAGCAAAAGCTCCATGATGTTGTT 720
QY 721 CTAGTGCACTGTTTGTGAGGATCTCCGCTCCGCAACATCCGTTATCCGTTATCATCAG 780
DB 1258 CTAGTGCACTGTTTGTGAGGATCTCCGCTCCGCAACATCCGTTATCCGTTATCATCAG 780
QY 781 AAGAGATGAGAAATGTTCTTTTGAATGAGCTTACAGATTTGTGAAGAAAGCACTACT 840
DB 1318 AAGAGATGAGAAATGTTCTTTTGAATGAGCTTACAGATTTGTGAAGAAAGCACTACT 840
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCCTGACATATGAAAGAAATTAAGAAC 900
DB 1378 ATATCTCCAACTTCAATTTTCTGGGCAACTCCTGACATATGAAAGAAATTAAGAAC 900
QY 901 CAGAGTGAAGATCAAGGCGCAAGAGCAAACTCAAGCTGTGCACTGGAAGAGCAAAAT 960
DB 1438 CAGAGTGAAGATCAAGGCGCAAGAGCAAACTCAAGCTGTGCACTGGAAGAGCAAAAT 960
QY 961 GAACCTGTCCCTGCTGTCTCAGAGGGTGAAGAAAGGAGACCGCTCTCAGTCCACC 1020
DB 1498 GAACCTGTCCCTGCTGTCTCAGAGGGTGAAGAAAGGAGACCGCTCTCAGTCCACC 1020
QY 1021 TGTGCGGACCTGTCTATCTCAGAGGCAAGAGCAAAAGCCGTCATCCGCGAGCGTG 1080
DB 1558 TGTGCGGACCTGTCTATCTCAGAGGCAAGAGCAAAAGCCGTCATCCGCGAGCGTG 1080
QY 1081 CCGAGGCTGCCAGCTGAGCGCTGAGCGTGTGTTAGAGGACACCGCTGTGTACAGCGCTG 1140
DB 1618 CCGAGGCTGCCAGCTGAGCGCTGAGCGTGTGTTAGAGGACACCGCTGTGTACAGCGCTG 1140
QY 1141 AGTGGGCTGCACTGTCTGCGCAGACAGCTGTGAAGACAGCAATTAAGCTCAAGCTTCT 1200
DB 1678 AGTGGGCTGCACTGTCTGCGCAGACAGCTGTGAAGACAGCAATTAAGCTCAAGCTTCT 1200
QY 1201 TCTTGATATCAATTCAGATTTCAATATTCAGCGAGCATGAGATCTTACATGAGCTTG 1260
DB 1738 TCTTGATATCAATTCAGATTTCAATATTCAGCGAGCATGAGATCTTACATGAGCTTG 1260
QY 1261 TCTTGATATCAATTCAGATTTCAATATTCAGCGAGCATGAGATCTTACATGAGCTTG 1260
DB 1798 TCTTGATATCAATTCAGATTTCAATATTCAGCGAGCATGAGATCTTACATGAGCTTG 1260
QY 1321 AAGCTATGCGAGTTCTCCCGTTGAGAGACTTATGAGAGCTGCGGAAACAGTCT 1380
DB 1858 AAGCTATGCGAGTTCTCCCGTTGAGAGACTTATGAGAGCTGCGGAAACAGTCT 1380
QY 1381 GATTAAGAGAGAGCGAGCATCCCGAAAGAGCTGACAGACCGGAGCTTCAAGAGCGAG 1440
DB 1918 GATTAAGAGAGAGCGAGCATCCCGAAAGAGCTGACAGACCGGAGCTTCAAGAGCGAG 1440
QY 1441 AGCAAGGATTTGATTCGATCGAGAAACAGAGAGAGTGGGACCGGCAAGAGTCTTTT 1500
DB 1978 AGCAAGGATTTGATTCGATCGAGAAACAGAGAGAGTGGGACCGGCAAGAGTCTTTT 1500
QY 1501 TCTCCACTGATGAAAGTGGAGCGTGAAGCAATTAACACCAAGTTGCTTTTGGG 1560
DB 2038 TCTCCACTGATGAAAGTGGAGCGTGAAGCAATTAACACCAAGTTGCTTTTGGG 1560
QY 2038 TCTCCACTGATGAAAGTGGAGCGTGAAGCAATTAACACCAAGTTGCTTTTGGG 1560
DB 1561 CTTTCCACAGCAGAGCACTCAAGAAAGTGTGTGGCTGTGGCTTTAAGGCTGGCAC 1620
QY 2098 CTTTCCACAGCAGAGCACTCAAGAAAGTGTGTGGCTGTGGCTTTAAGGCTGGCAC 1620
DB 1621 TCGATATTTTGGCCCTTCAAGCTTACCTTCCCTGACAGAGCTGTATTTTGGC 1680
QY 2158 TCGATATTTTGGCCCTTCAAGCTTACCTTCCCTGACAGAGCTGTATTTTGGC 1680
DB 1681 ACAGAGTCTCACTTACTCTGCTGAGCTTCAAGAGCACTGCAAGTACTCT 1740
```


QY 1141 AGTGGGCTGCACTGCTCCGACAGAGCTGGAGACAGCAATAGCTCAAGCTTCTTC 1200
DB 1678 AGTGGGCTGCACTGCTCCGACAGAGCTGGAGACAGCAATAGCTCAAGCTTCTTC 1737
QY 1201 TCTCTGATATCAATCAATGCTTCAATATTCAGCCGAGCATGAGACATCTTAACTGCTTC 1260
DB 1738 TCTCTGATATCAATCAATGCTTCAATATTCAGCCGAGCATGAGACATCTTAACTGCTTC 1797
QY 1261 TCTCTATCAAGAAAGTGGCTTGGAAATCTAATAAATCTTCACTAGCTGGAGGAGCAAC 1320
DB 1798 TCTCTATCAAGAAAGTGGCTTGGAAATCTAATAAATCTTCACTAGCTGGAGGAGCAAC 1857
QY 1321 AAGCTATGCGAGTCTCCCTCTGTTAGGAACTATGAGACAGTCCCGAAACCACTCTT 1380
DB 1858 AAGCTATGCGAGTCTCCCTCTGTTAGGAACTATGAGACAGTCCCGAAACCACTCTT 1917
QY 1381 GATTAAGAGAGAGCGAGCATCTCCGACAGAGCTGCAAGCCGAGGCTTTCAGACAGCAG 1440
DB 1918 GATTAAGAGAGAGCGAGCATCTCCGACAGAGCTGCAAGCCGAGGCTTTCAGACAGCAG 1977
QY 1441 AGCAAGCGATTGCAATCGGTCAAGAACAGCAGCAGTGGACCGGCGAGAGTCCCTTTA 1500
DB 1978 AGCAAGCGATTGCAATCGGTCAAGAACAGCAGCAGTGGACCGGCGAGAGTCCCTTTA 2037
QY 1501 TCTCCAGTGCATCGAAGGAGGAGGCTGGAGCAATTAACAACACAGCTTCTTTTCGAG 1560
DB 2038 TCTCCAGTGCATCGAAGGAGGAGGCTGGAGCAATTAACAACACAGCTTCTTTTCGAG 2097
QY 1561 CTTTTCACAGCAGCAGCAGCAGTCTGAGAGTCTGCTGGCTGGCTTAAAGGCTTGGCAG 1620
DB 2098 CTTTTCACAGCAGCAGCAGCAGTCTGAGAGTCTGCTGGCTGGCTTAAAGGCTTGGCAG 2157
QY 1621 TCGATATCTTGGGCCCCCAGACCTTACCTCCCTTCCCTGACAGCAGCTGTAATTTTTCG 1680
DB 2158 TCGATATCTTGGGCCCCCAGACCTTACCTCCCTTCCCTGACAGCAGCTGTAATTTTTCG 2217
QY 1681 ACAAGTCTTCACTTCTTACTGCTGCTGAGCCATCTAGAGAGGAGGCTGCTTACTCT 1740
DB 2218 ACAAGTCTTCACTTCTTACTGCTGCTGAGCCATCTAGAGAGGAGGCTGCTTACTCT 2277
QY 1741 GCCTACAGCTGAGCAGCAGCTGCCCACTTGGGAGACCAAGTCTATTTGTCGAGGCGG 1800
DB 2278 GCCTACAGCTGAGCAGCAGCTGCCCACTTGGGAGACCAAGTCTATTTGTCGAGGCGG 2337
QY 1801 CAGAAAGCAGATGACAGAGCTGCTGCGCGAGAGTGGCATGAGAGAGAGCCCTTTTGA 1860
DB 2338 CAGAAAGCAGATGACAGAGCTGCTGCGCGAGAGTGGCATGAGAGAGAGCCCTTTTGA 2397
QY 1861 AAGCAATTTAAAGCAGAGAGCTGCCAATGGAATTTGAGAGAGAGCATGTCAGAGAAC 1920
DB 2398 AAGCAATTTAAAGCAGAGAGCTGCCAATGGAATTTGAGAGAGAGCATGTCAGAGAAC 2457
QY 1921 AGGTCAAGGAGAGAGCTGGGAGAGTGGGAGTCACTAGCTTTTGGGAGAGTGA 1980
DB 2458 AGGTCAAGGAGAGAGCTGGGAGAGTGGGAGTCACTAGCTTTTGGGAGAGTGA 2517
QY 1981 ATCATTTAGAGTCTCC 1995
DB 2518 ATCATTTAGAGTCTCC 2532

RESULT 3
AX260342
LOCUS AX260342 1998 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 3 from Patent WO0173059.
ACCESSION AX260342
VERSION AX260342.1 GI:16509305
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1

AUTHORS Meyers, R.A.
TITLE 38692 and 21117: dual specificity phosphatase molecules and uses
JOURNAL Patent: WO 0173059-A 3 04-OCT-2001;
FEATURES Millennium Pharmaceuticals, Inc. (US)
SOURCE Location/Qualifiers
1. 1998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1991.8; DB 6; Length 1998;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCATGAGATGATTTGGAACCAATTGTTACTGAGAGTGGCTGCTGTGGA 60
DB 1 ATGGCCATGAGATGATTTGGAACCAATTGTTACTGAGAGTGGCTGCTGTGGA 60
QY 61 AGTGAACGGAAGAGTGTCTGCTAATTTGATGAGCCGCAATTTGTGAAATCAATATCC 120
DB 61 AGTGAACGGAAGAGTGTCTGCTAATTTGATGAGCCGCAATTTGTGAAATCAATATCC 120
QY 121 CACATTTTGAAGCCATTAATATCACTGCTCCAGCTTATGAGCGAAGTTCACACAG 180
DB 121 CACATTTTGAAGCCATTAATATCACTGCTCCAGCTTATGAGCGAAGTTCACACAG 180
QY 181 GACAAAGTGTAAATACAGAGCTATCCAGATTCAGGAAACATTAAGTTGACATTTGAT 240
DB 181 GACAAAGTGTAAATACAGAGCTATCCAGATTCAGGAAACATTAAGTTGACATTTGAT 240
QY 241 TGCAGTCAGAAAGTGTGATTTAGATGAAAGCTCCCAAGTGTGCTCTCTCTTCA 300
DB 241 TGCAGTCAGAAAGTGTGATTTAGATGAAAGCTCCCAAGTGTGCTCTCTCTTCA 300
QY 301 GACTGTTTCTACAGTACTTCTGAGTAACTGAGAAAGCTTCAACTCTGTCACCTG 360
DB 301 GACTGTTTCTACAGTACTTCTGAGTAACTGAGAAAGCTTCAACTCTGTCACCTG 360
QY 361 CTTCGAGGTTGGCTGAGTCTCTGCTGTTTCCCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 CTTCGAGGTTGGCTGAGTCTCTGCTGTTTCCCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 ACTGATGCTTCACTGATTTCTGAGCTTCTGCTTCACTGCTGCTGCTGCTGCTGCTG 480
DB 421 ACTGATGCTTCACTGATTTCTGAGCTTCTGCTTCACTGCTGCTGCTGCTGCTGCTG 480
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGAGAGTCTCTCAACAGAGCTGATA 540
DB 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGAGAGTCTCTCAACAGAGCTGATA 540
QY 541 CAGAGAAATGGAATTTGTTATGTTAAATGTCAGCTATACCTGTCGAAGCTGACTT 600
DB 541 CAGAGAAATGGAATTTGTTATGTTAAATGTCAGCTATACCTGTCGAAGCTGACTT 600
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 CCGTGTGTTGGAACAATCAGTATGTTCAATTTGAGAAAGCAAGAGCTCCCAATGATGTT 720
DB 661 CCGTGTGTTGGAACAATCAGTATGTTCAATTTGAGAAAGCAAGAGCTCCCAATGATGTT 720
QY 721 CTAGTCACTGTTTATGCTGAGTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 721 CTAGTCACTGTTTATGCTGAGTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 781 AAGAGAGTGAAGTCTTTAGATGAGAGCTTACAGATTTGTGAAAGAAAGAAAGAAAGCT 840
DB 781 AAGAGAGTGAAGTCTTTAGATGAGAGCTTACAGATTTGTGAAAGAAAGAAAGAAAGCT 840
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 900

Dp	841	ATATCTCCAACTTCAATTCTTTGTGGGCGAACTCTGGACTGTAGAGAAAGATTBAAGAC	900
Qy	901	CAGACTGGAGATCAGGGCCAAAGACAACTCAAGCTGCTGCACTTGGAGAGCCAAAT	960
Dp	901	CAGACTGGAGATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTTGGAGAGCCAAAT	960
Qy	961	GAACCTGTCCCTGCTGTCTCAGAGGGTGAACAGAAAAGCCAGAGCCCTTCATGTCACC	1020
Dp	961	GAACCTGTCCCTGCTGTCTCAGAGGGTGAACAGAAAACCGAGAGCCCTTCATGTCACC	1020
Qy	1021	TGTGGCGACTGTGCTAAGCTCAGAGGCAACAGAAAGAGCCCGTGCATCCCGCAGCGTG	1080
Dp	1021	TGTGGCGACTGTGCTAAGCTCAGAGGCAACAGAAAGAGCCCGTGCATCCCGCAGCGTG	1080
Qy	1081	CCGACGCTGCGCAGCGGTGACAGCCGTGCTGTAGAGACAGCCCGCTGTACAGCGCTC	1140
Dp	1081	CCGACGCTGCGCAGCGGTGACAGCCGTGCTGTAGAGACAGCCCGCTGTACAGCGCTC	1140
Qy	1141	AGTGGGCTGACCTGTGTCGACAGACAGGCTGAAAGACAGCAATTAAGCTCAAGCGTTCCT	1200
Dp	1141	AGTGGGCTGACCTGTGTCGACAGACAGGCTGAAAGACAGCAATTAAGCTCAAGCGTTCCT	1200
Qy	1201	TCTCTGGAATTCAAATTCAGTTTCATATTCAGCCAGCATGGCAGCATCTTACATGCTTC	1260
Dp	1201	TCTCTGGAATTCAAATTCAGTTTCATATTCAGCCAGCATGGCAGCATCTTACATGCTTC	1260
Qy	1261	TCCTCATCAGAAAGATGCTTTGGAAATACAAACCTTCACATCTGAGATGGGACCAAC	1320
Dp	1261	TCCTCATCAGAAAGATGCTTTGGAAATACAAACCTTCACATCTGAGATGGGACCAAC	1320
Qy	1321	AAGCTATGCGAGTTCTCCCTGTTCAGGAACATATCGAGACAGACTCCGAAACCAAGTCT	1380
Dp	1321	AAGCTATGCGAGTTCTCCCTGTTCAGGAACATATCGAGACAGACTCCGAAACCAAGTCT	1380
Qy	1381	GATTAAGAGAGAAAGCCAGCATCCCAAGAAAGCTGCGACCCGCAAGGCTTCACAGACCCAG	1440
Dp	1381	GATTAAGAGAGAAAGCCAGCATCCCAAGAAAGCTGCGACCCGCAAGGCTTCACAGACCCAG	1440
Qy	1441	AGCAAGCGATTTGCAATTCGGTGCAGAACCAAGCAGCAGTGGCACCGCCAGAGGTCCTTTTA	1500
Dp	1441	AGCAAGCGATTTGCAATTCGGTGCAGAACCAAGCAGCAGTGGCACCGCCAGAGGTCCTTTTA	1500
Qy	1501	TCTCCACTGTCATCGAAGTGGAGCGCTGAGAGCAATTACACACCAGCTTCCTTTTGGC	1560
Dp	1501	TCTCCACTGTCATCGAAGTGGAGCGCTGAGAGCAATTACACACCAGCTTCCTTTTGGC	1560
Qy	1561	CTTTTCACACAGCCAGACAGCACTTCAGAAATCTGTGGCCCTGAGGCTTGCGAC	1620
Dp	1561	CTTTTCACACAGCCAGACAGCACTTCAGAAATCTGTGGCCCTGAGGCTTGCGAC	1620
Qy	1621	TGCGATATCTTTGGCCCCCGACAACTCTACCCCTTCCCGTACAGAGCTGTGATTTTGGC	1680
Dp	1621	TGCGATATCTTTGGCCCCCGACAACTCTACCCCTTCCCGTACAGAGCTGTGATTTTGGC	1680
Qy	1681	ACAGAGTCTTCACATTTCTACTCTGCTCAGGCATCTACGAGAGCGAGTCCAGTTACTCT	1740
Dp	1681	ACAGAGTCTTCACATTTCTACTCTGCTCAGGCATCTACGAGAGCGAGTCCAGTTACTCT	1740
Qy	1741	GCTTACAGCTGACGCCAGCTGCGCATTTGCGAGACCAAGTCTATTTCTGTGCGACGCGG	1800
Dp	1741	GCTTACAGCTGACGCCAGCTGCGCATTTGCGAGACCAAGTCTATTTCTGTGCGACGCGG	1800
Qy	1801	CAGAAAGCCAAAGTGAACAGAGCTGACCTGCGGCGGAGCTGACATGAAGAGAGCCCTTTGAA	1860
Dp	1801	CAGAAAGCCAAAGTGAACAGAGCTGACCTGCGGCGGAGCTGACATGAAGAGAGCCCTTTGAA	1860
Qy	1861	AAGCAGTTTAAACGACAGAAAGCTGCCAAATGGAATTTGAGAGAGCATATGTCAGAGAAC	1920
Dp	1861	AAGCAGTTTAAACGACAGAAAGCTGCCAAATGGAATTTGAGAGAGCATATGTCAGAGAAC	1920
Qy	1921	AGGTCAAGGAAAGACTGAGGAGAAAGTGGACGTCAGTCTTACGCTTTTGGGACGACATGGA	1980

RESULT 4	AX278461	3059 bp	DNA	linear	PAT 02-NOV-2001
LOCUS	AX278461				
DEFINITION	Sequence 1 from Patent WO0177340.				
ACCESSION	AX278461				
VERSION	AX278461.1	GI:16605915			
KEYWORDS					
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
JOURNAL	Identification of a dual specificity phosphatase: dusp-10				
FEATURES	Patent: WO 0177340-A 1 18-OCT-2001;				
SOURCE	MERCK PATENT GmbH (DB)				
CDS	Location/Qualifiers				
	1..3059				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	127..2124				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD10545.1"				
	/db_xref="GI:16605916"				
	/translation="MAHEMIGTQIIVTERLIVALLLESGTEKYLIDSRPFVYNTSHILE				
	AININCKMKMRRLQODKVLITELIOLISAHKVIDDSQKVVYVDDSDVVALSPDC				
	FLTVLLGKLEKSRFNVHLNLAGRAFRRCPCGLCEGKSTVLPVTCISQPCIPVNIQGT				
	RLIPNLVLCQQRDLNKLKELQNGIIVGLVLAASNTCPKPDIPESHFLRVVNDSPCEK				
	ILPMLDKSVDFIERKAKASNCVIVHCLAGISRSATIAIAYIMKMDMSLDEAVRVEE				
	KRPITSNFNFILGQLDIYEKKIKNOTGASGPSKTLKILHEKPMRPPVAPEGQKSE				
	TPSPPCADATSEAGORPVHPASVPSVVOPLLEDSPLOVALSGHLISDRLED				
	SNLKRSFSLDIKSVSYSAASMAASHFSSSEDALEVKRSTLLDGNKLCSPPQVE				
	LSDEPTSPDKKEASTIPKQLQTPRPDSQSKHLVRISSSGTAQSRSLPLRRSS				
	VENDYATPSLFGISTSQLTKSAGJLKGMSHDILAPQTSPLSTVSWFATESSHF				
	YSSAIIYGSASYSAYSCSQLPTCGQVYVVRARQKSDRADRSRWHEBSPEKQPK				
	RRSQMRERGSIMSERNRRELGVGSQSSFSGSMETIEVVS"				
ORIGIN					
Query Match	99.8%; Score 1991.8; DB 6; Length 3059;				
Best Local Similarity	99.9%; Pred. No. 0;				
Matches 1993; Conservative	0; Mismatches 2; Indels 0; Gaps 0;				
QY	1 ATGCCCCATGATGATTTGGAACTCAATTTGTTACTGAGAGCTTGCTGCTCTGCTGAA	60			
DB	127 ATGCCCCATGAGATGATGGAACTCAATTTGTTACTGAGAGCTTGCTGCTGCTGAA	186			
QY	61 AGTGAACGGAAGAAAGTCTGCTTAATTGATAGCGCGCATTTTGGAATACAATACATCC	120			
DB	187 AGTGAACGGAAGAAAGTCTGCTTAATTGATAGCGCGCATTTTGGAATACAATACATCC	246			
QY	121 CACATTTTGAAGCCATTAAATATCACTGCTCCACAGCTTATGAAGCGAAGTTGCAACAG	180			
DB	247 CACATTTTGAAGCCATTAAATATCACTGCTCCACAGCTTATGAAGCGAAGTTGCAACAG	306			
QY	181 GACAAAGTGTAAATTAACAGAGCTCATTCACATTCAGCGGAAACATAAGTGTGACATTGAT	240			
DB	307 GACAAAGTGTAAATTAACAGAGCTCATTCACACATTACGGAACATAAGTGTGACATTGAT	366			
QY	241 TGCAAGTCAAGAGTTGAGTTTACGATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCA	300			
DB	367 TGCAAGTCAAGAGTTGAGTTTACGATCAAGAGCTCCCAAGATGTTGCTCTCTCTCTTCA	426			
QY	301 GACTGTTTCTCACTGACTCTTGAGTAACTGAGAAAGCTTCAACTCTGTTACCTG	360			

Db 427 GACTGTTTCTGCTGTACTTCTGGTAACTGGAGAAAGACTTCAACTGTTCACTG 486
Qy 361 CTGGAGAGGAGGTTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
Db 487 CTGGAGAGGAGGTTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 546
Qy 421 ACTGATGCTTCCCTGAGATTTCTCAGAGCTTGTACTGCTGTGCAACATTGGGCAAC 480
Db 547 ACTGATGCTTCCCTGAGATTTCTCAGAGCTTGTACTGCTGTGCAACATTGGGCAAC 606
Qy 481 CGAATTCTTCCCAATCTTATCTTGTGCTGCCAGAGATGTCTCTCAACAAGAGCGATTA 540
Db 607 CGAATTCTTCCCAATCTTATCTTGTGCTGCCAGAGATGTCTCTCAACAAGAGCGATTA 666
Qy 541 CAGCAAAATGGGAATGGTTATGTGTAAATGGCAGCTAATACCTGTCAAGCTGACTTT 600
Db 667 CAGCAAAATGGGAATGGTTATGTGTAAATGGCAGCTAATACCTGTCAAGCTGACTTT 726
Qy 601 ATCCCGAGTCTCAATTTCCCTGCTGCTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 660
Db 727 ATCCCGAGTCTCAATTTCCCTGCTGCTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 786
Qy 661 CCGTGTGGGCAAAATCAGTATTTCAATTTGAGAAAGAAAAGCTTCAATGATGTT 720
Db 787 CCGTGTGGGCAAAATCAGTATTTCAATTTGAGAAAGAAAAGCTTCAATGATGTT 846
Qy 721 CTAGTCACTGTTTACTGGGATCTCCGCTCCGCTCCGACATGCTATGCTATCATGATG 780
Db 847 CTAGTCACTGTTTACTGGGATCTCCGCTCCGCTCCGACATGCTATGCTATCATGATG 906
Qy 781 AAGAGATGAGATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 907 AAGAGATGAGATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 966
Qy 841 ATATCTCCAACTTCAATTTCTGGGCAACTCTGTGACTATGAGAAAGAAATTAAGAAC 900
Db 967 ATATCTCCAACTTCAATTTCTGGGCAACTCTGTGACTATGAGAAAGAAATTAAGAAC 1026
Qy 901 CAGACTGAGATCAGAGGCGCAAGAGCAAACTCAGCTGTGCACTGGAGAAAGCAAT 960
Db 1027 CAGACTGAGATCAGAGGCGCAAGAGCAAACTCAGCTGTGCACTGGAGAAAGCAAT 1086
Qy 961 GAACCTGTCCCTGCTGTCTCAGAGGGGTGAGCAGAAAAGCGAAGCGCCCTCAGTCAACC 1020
Db 1087 GAACCTGTCCCTGCTGTCTCAGAGGGGTGAGCAGAAAAGCGAAGCGCCCTCAGTCAACC 1146
Qy 1021 TGTGCGCACTCTGCTGACTCAGAGGCGAGAGCAAAAGCCCGTGCATCCCGCAAGCTG 1080
Db 1147 TGTGCGCACTCTGCTGACTCAGAGGCGAGAGCAAAAGCCCGTGCATCCCGCAAGCTG 1206
Qy 1081 CCCAGCGTCCGAGCGTGCAGCGCTGCTGTGAGAGCAGCCCGCTGTGTACAGCGCTG 1140
Db 1207 CCCAGCGTCCGAGCGTGCAGCGCTGCTGTGAGAGCAGCCCGCTGTGTACAGCGCTG 1266
Qy 1141 AGTGGGCTCAGCTGTGCGCAGACAGGCTGGAAGACAGCAATAGCTCAAGCCCTTCTTC 1200
Db 1267 AGTGGGCTCAGCTGTGCGCAGACAGGCTGGAAGACAGCAATAGCTCAAGCCCTTCTTC 1326
Qy 1201 TCTCTGATATCAATCAGTTTCAATTTCAAGCGAGCTGAGAGATCTTTCATGAGCTTTC 1260
Db 1327 TCTCTGATATCAATCAGTTTCAATTTCAAGCGAGCTGAGAGATCTTTCATGAGCTTTC 1386
Qy 1261 TCCATATCAAGAAATGCTTTTGAATTAACAACCTTCACTCTGATGGAGCAAC 1320
Db 1387 TCCATATCAAGAAATGCTTTTGAATTAACAACCTTCACTCTGATGGAGCAAC 1446
Qy 1321 AAGCTATGCGAGTTCTCCCTGTTTCAAGAACTATCGAGCAAGTCTCCGAAACAGTCTCT 1380
Db 1447 AAGCTATGCGAGTTCTCCCTGTTTCAAGAACTATCGAGCAAGTCTCCGAAACAGTCTCT 1506
Qy 1381 GATTAAGAGAGAGCAGATCCCAAGAACTGCAAGCCGCAAGGCTTTCAGACAGGCTG 1440

Db 1507 GATTAAGAGAGAGCAGATCCCAAGAAAGCTGACAGACCGCCAGGCTTTCAGACAGCAG 1566
Qy 1441 AGCAAGGAGATGATTCCTGCTCAGAACCGACAGAGTGGCAACCGCCAGAGGTCCTTTTA 1500
Db 1567 AGCAAGGAGATGATTCCTGCTCAGAACCGACAGAGTGGCAACCGCCAGAGGTCCTTTTA 1626
Qy 1501 TCTTCCAGCTGATCAGAGTGGAGAGCGTGGAGAGCAATTAACAACAGCTTCTTTTGGG 1560
Db 1627 TCTTCCAGCTGATCAGAGTGGAGAGCGTGGAGAGCAATTAACAACAGCTTCTTTTGGG 1686
Qy 1561 CTTTCCACAGCAGACAGACACTTCAAGAGTCTGTGCTGCTGAGCTTAAAGGCTGGCAC 1620
Db 1687 CTTTCCACAGCAGACAGACACTTCAAGAGTCTGTGCTGCTGAGCTTAAAGGCTGGCAC 1746
Qy 1621 TCGGATATCTTGGGCCCCCAGACCTTCAACCTTCCCTGACAGAGCTGTGATTTTGGC 1680
Db 1747 TCGGATATCTTGGGCCCCCAGACCTTCAACCTTCCCTGACAGAGCTGTGATTTTGGC 1806
Qy 1681 ACAGAGTCTTCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1807 ACAGAGTCTTCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1866
Qy 1741 GCTTACGCTGCAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db 1867 GCTTACGCTGCAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1926
Qy 1801 CAGAGCAGAGTGAAGAGCTGATCTGCGGCGGAGAGCTGAGTGAAGAGAGCCCTTTGAA 1860
Db 1927 CAGAGCAGAGTGAAGAGCTGATCTGCGGCGGAGAGCTGAGTGAAGAGAGCCCTTTGAA 1986
Qy 1861 AAGCAGTTTAAACGAGAGAGCTGCAATGGAATTTGAGAGAGCATCATGTAGAGAAC 1920
Db 1987 AAGCAGTTTAAACGAGAGAGCTGCAATGGAATTTGAGAGAGCATCATGTAGAGAAC 2046
Qy 1921 AGGTCAAGGAGAGAGCTGAGGAGAGTGGGAGTCAAGTCTGCTTTTGGGCAAGTGA 1980
Db 2047 AGGTCAAGGAGAGAGCTGAGGAGAGTGGGAGTCAAGTCTGCTTTTGGGCAAGTGA 2106
Qy 1981 ATCATTTGAGGTCTCC 1995
Db 2107 ATCATTTGAGGTCTCC 2121

RESULT 5
AX441210 3496 bp DNA linear PAT 28-JUN-2002
LOCUS Sequence 1 from Patent WO0226997.
DEFINITION AX441210
ACCESSION AX441210
VERSION AX441210.1 GI:21665766
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Luche, R. M. and Wei, B.
AUTHORS Dep-16 dual-specificity phosphatase
TITLE Patent: WO 0226997-A 1 04-APR-2002;
JOURNAL Ceptlyr, Inc. (US)
FEATURES
source
1. 3496
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1991.8; DB 6; Length 3496;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCCCATGAGATGATTTGGAATCAATTTTACTGAGAGGTTGGTGTGCTGAGAA 60
Db 562 ATGGCCCATGAGATGATTTGGAATCAATTTTACTGAGAGGTTGGTGTGCTGAGAA 621

QY 61 AGTGAACGAAAAAGCTGCTTAATGATAGCCGGCATTTGTGAAATACATATCC 120
DB 622 AGTGAACGAAAAAGCTGCTTAATGATAGCCGGCATTTGTGAAATACATATCC 681
QY 121 CACATTTTGAAGCATTATATCACTGCTCCAGCTTATGAAGGAAAGTTGCAACAG 180
DB 682 CACATTTTGAAGCATTATATCACTGCTCCAGCTTATGAAGGAAAGTTGCAACAG 741
QY 181 GACAAAGTTTAAATTACAGAGCTCATCAGCAATTCAGGAAACATAGTTGACATTGAT 240
DB 742 GACAAAGTTTAAATTACAGAGCTCATCAGCAATTCAGGAAACATAGTTGACATTGAT 801
QY 241 TGCAGTCAGAAAGTTGATGTTATGATCAAGCAAGCTCCCAAGATGTGCTCTCTTCA 300
DB 802 TGCAGTCAGAAAGTTGATGTTATGATCAAGCAAGCTCCCAAGATGTGCTCTCTTCA 861
QY 301 GACTGTTTCTCACTGACTTCTGAGTAACTGGAGAAAGACTTCAACTCTGTTCACTG 360
DB 862 GACTGTTTCTCACTGACTTCTGAGTAACTGGAGAAAGACTTCAACTCTGTTCACTG 921
QY 361 CTTCGAGAGTGGTTTGTGCTGAGTCTCTGTTGTTCCCTGAGGAAAGAAATCC 420
DB 922 CTTCGAGAGTGGTTTGTGCTGAGTCTCTGTTGTTCCCTGAGGAAAGAAATCC 981
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTTACCTGTTGCAATTTGGCCCAAC 480
DB 982 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTTACCTGTTGCAATTTGGCCCAAC 1041
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGAGGAGATGCTCCCAAGAGAGCTGATA 540
DB 1042 CGAATTTCTCCCAATCTTTATCTTGGCTGAGGAGATGCTCCCAAGAGAGCTGATA 1101
QY 541 CAGCAGAAATGGATTTGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 600
DB 1102 CAGCAGAAATGGATTTGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 1161
QY 601 ATCCCGCAGTCTCATTTCTGCTGCTGCTGCTGATGAATGACGCTTTTGTGAAATTTTG 660
DB 1162 ATCCCGCAGTCTCATTTCTGCTGCTGCTGCTGATGAATGACGCTTTTGTGAAATTTTG 1221
QY 661 CCGGTTTGGCAAAATGATGATTTTCAATGAGAAAGCAAAAGCCTCAATGATGTT 720
DB 1222 CCGGTTTGGCAAAATGATGATTTTCAATGAGAAAGCAAAAGCCTCAATGATGTT 1281
QY 721 CTAGTGCACTGTTTGTGCTGGATCTCCGCTCCGCAATCGCTATCGCTACATCATG 780
DB 1282 CTAGTGCACTGTTTGTGCTGGATCTCCGCTCCGCAATCGCTATCGCTACATCATG 1341
QY 781 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTACT 840
DB 1342 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTACT 1401
QY 841 ATATCTCAAACTCAATTTTCTGGGCAACTCCTGACTATGAGAGAAATTAAGAAC 900
DB 1402 ATATCTCAAACTCAATTTTCTGGGCAACTCCTGACTATGAGAGAAATTAAGAAC 1461
QY 901 CAGAGTGAAGATCAGGAGCCAAAGAGCAAACTCAAGTGTGCTGCACTGAGAAAGCCAAAT 960
DB 1462 CAGAGTGAAGATCAGGAGCCAAAGAGCAAACTCAAGTGTGCTGCACTGAGAAAGCCAAAT 1521
QY 961 GAACCTGTCCCTGTCTGTCTAGAGAGGTGAGACAAAGAGCCGCTCACTCCACC 1020
DB 1522 GAACCTGTCCCTGTCTGTCTAGAGAGGTGAGACAAAGAGCCGCTCACTCCACC 1581
QY 1021 TGAGCAGACTGTGCTTACTCTGAGAGCAGAGCAAAAGCCCGGTCATCCGACAGGTG 1080
DB 1582 TGAGCAGACTGTGCTTACTCTGAGAGCAGAGCAAAAGCCCGGTCATCCGACAGGTG 1641
QY 1081 CCCAGCGTCCAGCAGCTGAGCCGCTGCTGTTAGAGACAGCCGCTGATACAGGCGCTC 1140
DB 1642 CCCAGCGTCCAGCAGCTGAGCCGCTGCTGTTAGAGACAGCCGCTGATACAGGCGCTC 1701

QY 1141 AGTGGCTGCACTCTGCTCCGCAAGACAGCTGAAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
DB 1702 AGTGGCTGCACTCTGCTCCGCAAGACAGCTGAAAGACAGCAATTAAGCTCAAGCTTCTTC 1761
QY 1201 TCTCTGATATCAATCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCA 1260
DB 1762 TCTCTGATATCAATCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCA 1821
QY 1261 TCTCTATCAAGAAATGCTTTGAAATCTAATAAATCTTCAATCTTCTGATGAGGACCAAC 1320
DB 1822 TCTCTATCAAGAAATGCTTTGAAATCTAATAAATCTTCAATCTTCTGATGAGGACCAAC 1881
QY 1321 AAGCTATGCAAGTTCTCCCTGTTTCAAGAACTATGAGAGCAAGCTCCCAACCAATGCT 1380
DB 1882 AAGCTATGCAAGTTCTCCCTGTTTCAAGAACTATGAGAGCAAGCTCCCAACCAATGCT 1941
QY 1381 GATTAAGAGAGAAAGCCAGATCCCAAGAAAGCTGAGACAGCCGCTTACAGACCCAG 1440
DB 1942 GATTAAGAGAGAAAGCCAGATCCCAAGAAAGCTGAGACAGCCGCTTACAGACCCAG 2001
QY 1441 AGCAAGCAATTTGATTTGCTGCTGAGAACACAGACAGAGTCCGACAGAGTCCCTTTTA 1500
DB 2002 AGCAAGCAATTTGATTTGCTGCTGAGAACACAGACAGAGTCCGACAGAGTCCCTTTTA 2061
QY 1501 TCTTCACTGATGAGAGTGGAGAGCTGAGAGCAATTTACCAACAGACTTCTTTTGGC 1560
DB 2062 TCTTCACTGATGAGAGTGGAGAGCTGAGAGCAATTTACCAACAGACTTCTTTTGGC 2121
QY 1561 CTTTTCAACAGCAGCAGACCTTCAAGAAATCTGCTGAGCTGAGCTTAAAGGCTGGCAC 1620
DB 2122 CTTTTCAACAGCAGCAGACCTTCAAGAAATCTGCTGAGCTGAGCTTAAAGGCTGGCAC 2181
QY 1621 TCGGATATCTTGGGCCCCCAGACCTTACCTTCCCTGACACAGAGCTGATTTTGGC 1680
DB 2182 TCGGATATCTTGGGCCCCCAGACCTTACCTTCCCTGACACAGAGCTGATTTTGGC 2241
QY 1681 ACAGATGCTCACTTCTACTGCTGCTGAGCCTTCAAGAGAGAGCTGCAATTTACTCT 1740
DB 2242 ACAGATGCTCACTTCTACTGCTGCTGAGCCTTCAAGAGAGAGCTGCAATTTACTCT 2301
QY 1741 GCCTACAGCTGACAGCAGCTGCCACTTGGAGAGCAAGCTTATTTCTGCTGAGGCGG 1800
DB 2302 GCCTACAGCTGACAGCAGCTGCCACTTGGAGAGCAAGCTTATTTCTGCTGAGGCGG 2361
QY 1801 CAGAAAGCAAGTACAGAGCTGACTGCTGCGGCGGAGCTGCTGATGAAGAGAGCCCTTTGAA 1860
DB 2362 CAGAAAGCAAGTACAGAGCTGACTGCTGCGGCGGAGCTGCTGATGAAGAGAGCCCTTTGAA 2421
QY 1861 AAGCAGTTTAAAGCAGAAAGCTGCAATGAAATTTGAGAGAGCATGTCAGAGAAC 1920
DB 2422 AAGCAGTTTAAAGCAGAAAGCTGCAATGAAATTTGAGAGAGCATGTCAGAGAAC 2481
QY 1921 AGCTACAGGAGAGAGCTGGGAGAAAGTGGAGAGTCACTTGTGAGGAGCATGAA 1980
DB 2482 AGCTACAGGAGAGAGCTGGGAGAAAGTGGAGAGTCACTTGTGAGGAGCATGAA 2541
QY 1981 ATCATTTAGGTCTCC 1995
DB 2542 ATCATTTAGGTCTCC 2556

RESULT 6
AB052156 3521 bp mRNA linear PRI 18-OCT-2001
LOCUS
DEFINITION Homo sapiens MKP-7 mRNA for MAPK phosphatase-7, complete cds.
ACCESSION AB052156
VERSION AB052156.1 GI:13548676
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1

AUTHORS Masuda, K., Shima, H., Maranabe, M. and Kikuchi, K.
TITLE MKP-7, a novel mitogen-activated protein kinase phosphatase,
functions as a shuttle protein
JOURNAL J. Biol. Chem. 276 (42), 39002-39011 (2001)
MEDLINE 2146429
PUBMED 11489891
REFERENCE 2 (bases 1 to 3521)
AUTHORS Masuda, K., Shima, H. and Kikuchi, K.
TITLE Direct Submision
Submitted (05-DEC-2000) Kouhei Masuda, Institute for Genetic
Medicine, Division of Biochemical Oncology and Immunology, Kita-ku
Kita-5-jou nishi7-tyoume, Sapporo, Hokkaido 060-0815, Japan
(E-mail:kouhei@hokudai.ac.jp, Tel.81-11-706-5536,
Fax:81-11-707-6839)

FEATURES

source Location/Qualifiers
1..3521
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/cell_line="U937"
1..3521
/gene="MKP-7"
564..2561
/gene="MKP-7"
/codon_start=1
/product="MAPK phosphatase-7"
/protein_id="BAB40814.1"
/db_xref="GI:13548677"
/translation="MAHEMIGTQIVTERLVALLESGTEKYLIDSRPFVYNTSHILE
AININGSKLMKRLQODKEMITLTIHSAKHVDIDSOXVVVVVSSDVASLSDC
RLPVLGLKLEKSFNSVHLGAFEPFRCPEGLCEKSTLVPICISQPLPVANGPT
ILPWLKSVDFLEKAKANGCVLHCHAGISRAITAIYIMKRDMSIDEXVRYEKE
KRPTSPFNFLGOLLDEYKIKKQTSAGSPSKLILHLEKNEPVPVASEGQKSE
TLPSPCADSATSEASARVHPVAPSVPSVQSLIEDSPVVALSGHLISADLIED
SNLKRSPSLDIKSVTSASMAASLHGFSSSEDALEYRSTTLGDTNKLQDFSVDE
LSROTPEPSDPKEASIPKQLQTPARPSDSQSKLHSHVRTSSSGTQORSLPIRHSVS
VEDNHTSPFLGSLTSQOHLTKSAGLGLKMGHSDIAPOTISPTSSMYRATSSHP
YSASATVIGSASVAYSCQLPTGPOVYSVRROKPRAPRASRSMHESPFEROKF
RRCQMFESGIMSINRSRELKVGSGSPSGSEITFVS"

ORIGIN

Query Match 99.8%; Score 1991.8; DB 9; Length 3521;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATGATGAACTCAATTTGTAAGAGAGTGGTGGCTCTCTGGAA 60
Db 564 ATGGCCCATGATGATGATGAACTCAATTTGTAAGAGAGTGGTGGCTCTCTGGAA 60
QY 61 AGTGAACCGAAGAAAGTGTGCTAATGATAGCCGGCAATTTGTGAATCAATATCATCC 120
Db 624 AGTGAACCGAAGAAAGTGTGCTAATGATAGCCGGCAATTTGTGAATCAATATCATCC 120
QY 121 CACATTTTGAAGCAATTAATCAACTGCTCAAGCTTATGAAGGGAAGTTTGCACAG 180
Db 684 CACATTTTGAAGCAATTAATCAACTGCTCAAGCTTATGAAGGGAAGTTTGCACAG 180
QY 181 GACAAAGTGTAAATTACAGAGCTCATCCAGATTCAAGCAATTAAGTTGACATTGAT 240
Db 744 GACAAAGTGTAAATTACAGAGCTCATCCAGATTCAAGCAATTAAGTTGACATTGAT 240
QY 241 TGCACTGCAAGGTTGATGATGATCAAGCTCCCAAGATTGCGCTCTCTCTTCA 300
Db 804 TGCACTGCAAGGTTGATGATGATCAAGCTCCCAAGATTGCGCTCTCTCTTCA 300
QY 301 GACTGTTTTCATCTGATCTTGGGGTAACTGGAAGAAGCTTCAACTCTGTCACTG 360
Db 864 GACTGTTTTCATCTGATCTTGGGGTAACTGGAAGAAGCTTCAACTCTGTCACTG 360
QY 361 CTTCAGAGTGGTGGTGGTCTCTGTTGTTTCCCTGGCTGTGAAGGAAATCC 420
Db 420 CTTCAGAGTGGTGGTGGTCTCTGTTGTTTCCCTGGCTGTGAAGGAAATCC 420

Db 924 CTTCAGAGTGGTGGTGGTCTCTGTTGTTTCCCTGGCTGTGAAGGAAATCC 983
QY 421 ACTTAGTCCCTACCGCATTTCTCAGAGCTTGTCTTACTGTGGCAACATGGGCCAAC 480
Db 984 ACTTAGTCCCTACCGCATTTCTCAGAGCTTGTCTTACTGTGGCAACATGGGCCAAC 480
QY 481 CGAATTTCTCCAAATCTTAACTTGGCTGCGCAGAGATGCTCCCAAGAGCTGATA 540
Db 1044 CGAATTTCTCCAAATCTTAACTTGGCTGCGCAGAGATGCTCCCAAGAGCTGATA 540
QY 541 CAGCAAAATGGGATTTGTTAATGTTAAAGCCAGATTAATCTGTCCAAAGCTTCACT 600
Db 1104 CAGCAAAATGGGATTTGTTAATGTTAAAGCCAGATTAATCTGTCCAAAGCTTCACT 600
QY 601 ATCCCGAGTCTCAATTTCTCGGTGTCCTGTGAATGACAGCTTTGTGAAGAAATTTTG 1163
Db 1164 ATCCCGAGTCTCAATTTCTCGGTGTCCTGTGAATGACAGCTTTGTGAAGAAATTTTG 1163
QY 661 CCGTGTGGTGAACAATCAGATGATTTCAATTGAGAAAGCAAAAGCTCCAAATGATGTT 720
Db 1224 CCGTGTGGTGAACAATCAGATGATTTCAATTGAGAAAGCAAAAGCTCCAAATGATGTT 720
QY 721 CTAGTGACCTTTAGCTGGGATCTCCCGCTCCGACCATGCTATGCGCTACATGATG 780
Db 1284 CTAGTGACCTTTAGCTGGGATCTCCCGCTCCGACCATGCTATGCGCTACATGATG 780
QY 781 AAGAGATGACATGCTTTAGTGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
Db 1344 AAGAGATGACATGCTTTAGTGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
QY 841 ATATCTCCAAATCTTCAATTTTGTGGGCCAACTCTGTGACTATGAGAAAGATTAGAAC 1403
Db 1404 ATATCTCCAAATCTTCAATTTTGTGGGCCAACTCTGTGACTATGAGAAAGATTAGAAC 1403
QY 901 CACATGAGCATCAAGGCGCAAGAGCAAACTCAAGCTGTGCACTGTGGAAGCCAAAT 960
Db 1464 CACATGAGCATCAAGGCGCAAGAGCAAACTCAAGCTGTGCACTGTGGAAGCCAAAT 960
QY 961 GAACCTGTCCCTGCTGTCTCAGAGGTGGAACAGAAAGCGCCCTCAGTCCAGCC 1020
Db 1524 GAACCTGTCCCTGCTGTCTCAGAGGTGGAACAGAAAGCGCCCTCAGTCCAGCC 1020
QY 1021 TGTGCGCACTGTGCTACTCAGAGGAGCAGAGCAAAAGCCGTGATCCGCCAGCTG 1080
Db 1584 TGTGCGCACTGTGCTACTCAGAGGAGCAGAGCAAAAGCCGTGATCCGCCAGCTG 1080
QY 1081 CCGAGCGTCCCGACCGCTGACCGCTGTGTTAAGAGCAGCCCGCTGTACAGGCGTTC 1643
Db 1644 CCGAGCGTCCCGACCGCTGACCGCTGTGTTAAGAGCAGCCCGCTGTACAGGCGTTC 1643
QY 1141 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1704 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
QY 1201 TCTCTGGAATGAATCAATCAGTTTCAATTCAGCCAGATGAGAGATCTTAAAGGCTTC 1260
Db 1764 TCTCTGGAATGAATCAATCAGTTTCAATTCAGCCAGATGAGAGATCTTAAAGGCTTC 1260
QY 1261 TCTCTATCAAGAGATGCTTGAATTAATCAAACTTCAATATCTGTGAAGGAGCAAC 1320
Db 1824 TCTCTATCAAGAGATGCTTGAATTAATCAAACTTCAATATCTGTGAAGGAGCAAC 1320
QY 1321 AAGCTATGCAATTTCCCTGTTTACAGAACTATCGAGAGACCTCCGAAACAGATCT 1380
Db 1884 AAGCTATGCAATTTCCCTGTTTACAGAACTATCGAGAGACCTCCGAAACAGATCT 1380
QY 1381 GATTAAGAGAGAGCCAGATCCCAAGAGTGTGAGACCGCGCTTCAAGACCCAG 1440
Db 1944 GATTAAGAGAGAGCCAGATCCCAAGAGTGTGAGACCGCGCTTCAAGACCCAG 1440
QY 1441 AGCAAGGATTTGCAATTTGGTTCAGAACTGAGAGAGCCGCCAGAGATCTCTTTTA 1500
Db 2004 AGCAAGGATTTGCAATTTGGTTCAGAACTGAGAGAGCCGCCAGAGATCTCTTTTA 2063

QY 1501 TCTCCATGCGATGGAAGTGGAGCGTGGAGAGCAATTACCAACAGCTTCTTTTCCGC 1560
Db 2064 TCTCCATGCGATGGAAGTGGAGCGTGGAGAGCAATTACCAACAGCTTCTTTTCCGC 2123
QY 1561 CTTTCCACGACGACGAGACGACCTCAGCAAGTCTGTGGCTTGGGCTTAAAGGCTGGAC 1620
Db 2124 CTTTCCACGACGACGAGACGACCTCAGCAAGTCTGTGGCTTGGGCTTAAAGGCTGGAC 2183
QY 1621 TCGGATATCTTGGGCCCCCAGACCTCTACCCCTTCCCTGACCGACGCTGTATTTTGGC 1680
Db 2184 TCGGATATCTTGGGCCCCCAGACCTCTACCCCTTCCCTGACCGACGCTGTATTTTGGC 2243
QY 1681 ACAGAGCTCTCAACTTCTACTCTGCTGACGATCTGAGGAGGAGGCAATGCAATTTCTCT 1740
Db 2244 ACAGAGCTCTCAACTTCTACTCTGCTGACGATCTGAGGAGGAGGCAATGCAATTTCTCT 2303
QY 1741 GCCTACAGCTGACGACGACCTGCTGCTGACGACCAAGTCTATTTCTGTGCGCAGGCG 1800
Db 2304 GCCTACAGCTGACGACGACCTGCTGCTGACGACCAAGTCTATTTCTGTGCGCAGGCG 2363
QY 1801 CAGAGCCAAAGTACAGAGCTGACTGCGGCGGAGAGCTGCGATGAAAGAGAGCCCTTTGAA 1860
Db 2364 CAGAGCCAAAGTACAGAGCTGACTGCGGCGGAGAGCTGCGATGAAAGAGAGCCCTTTGAA 2423
QY 1861 AAGCAGTTTAAAGCGCAAGCTGCAATGGAATTTGAGAGAGCATCATGTGAGAGAAC 1920
Db 2424 AAGCAGTTTAAAGCGCAAGCTGCAATGGAATTTGAGAGAGCATCATGTGAGAGAAC 2483
QY 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGAGAGTCAAGTCTGCTTTCGGGACAGATGAA 1980
Db 2484 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGAGAGTCAAGTCTGCTTTCGGGACAGATGAA 2543
QY 1981 ATCATTTAGAGTCTCC 1995
Db 2544 ATCATTTAGAGTCTCC 2558

RESULT 7
AX260340 3544 bp DNA linear PAT 26-OCT-2001
LOCUS Sequence 1 from Patent WO0173059.
DEFINITION AX260340
ACCESSION AX260340
VERSION AX260340.1 GI:16509303
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Meyers, R.A.
TITLES 38692 and 21117: dual specificity phosphatase molecules and uses
JOURNAL Patent: WO 0173059-A 1 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Source Location/Qualifiers
1. 3544
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
589..2586
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10218.1"
/db_xref="GI:16509304"
/db_xref="REMBTREMBL:CAD10218"
/translation="MAHBMGTQIIVTRLVALLSGNKKVLLIDSRPVEVYNTSHLE
AIVNISKMRRLQDDKVLITELIHSAGKVDIDCSQVYVYVSSDDVLSDDC
FLVVLGKLEKSTNSVHLAAGFAFBSRCPGLCEGKSTLVPCISQPLPANTPT
RILPNDLYGQDVLNKEMLQNGIGVLAASNTCPDPFIPESHFLRVAVNDFCEK
ILPMLDSVPIEKAKNSGCVLVHICAGISRSATIAIAYIMRMDSIDEAAYRVEK
KRPTISPENFPLGQLDYEKIKNGTGASGPKSKLILHEKNEBPVAVSEGGQGS
TPISPCADSATSEAAAGRPVHPASVPSVSVQPSLLESDPLVQALSGHLSDRLSD

SNKLRFSFLIDIKSVYSASMAASHGSSSEDALEYKPSYTLIDGTNLCQPSPVQE
LSEORPENSPEDEBASIPKLIOTAPSPDSORLHSVTRSSGTAQORSLISPLHRSQS
VEDNHTSRLGRLSTSOHLTSGAGLGKMSDILAPOTSPSLTSSVYFATHSRHF
YASAIYIGSASYSAYSQSGQLPTCCDQVYSVRROKPSRABRSRMSHESPEKQPK
RRSCMERSESTMSRREELKGVSSQSSFSGSMELIIVS"

Query Match 99.8%; Score 1991.8; DB 6; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGAACTCAATTGTTACTGAGAGTGGCTGCTGCGAA 60
Db 589 ATGGCCCATGAGATGATGAACTCAATTGTTACTGAGAGTGGCTGCTGCGAA 648
QY 61 AGTGAACGAAAAAGTCTGCTAATTTGATAGCCGCCCATTTGTGGAATCAATATCC 120
Db 649 AGTGAACGAAAAAGTCTGCTAATTTGATAGCCGCCCATTTGTGGAATCAATATCC 708
QY 121 CACATTTGGAAAGCCATTAAATCAACGCTCCAGCTTATGAAAGGAAAGTTGCAACAG 180
Db 709 CACATTTGGAAAGCCATTAAATCAACGCTCCAGCTTATGAAAGGAAAGTTGCAACAG 768
QY 181 GACAAAGTGTAAATACAGAGCTCATCCAGATTCAGCGAAACATPAGGTTGACATTTAT 240
Db 769 GACAAAGTGTAAATACAGAGCTCATCCAGATTCAGCGAAACATPAGGTTGACATTTAT 828
QY 241 TGCAGTCAGAGGTTGATTTACATCAAGTCCCAAGATGTTGCTCTCTTTCA 300
Db 829 TGCAGTCAGAGGTTGATTTACATCAAGTCCCAAGATGTTGCTCTCTTTCA 888
QY 301 GACTGTTTCTCACTGATTTCTGGGTAACTGAGAAAGAGCTCACTCTGTCACCTG 360
Db 889 GACTGTTTCTCACTGATTTCTGGGTAACTGAGAAAGAGCTCACTCTGTCACCTG 948
QY 361 CTTCGAGTGGGTTGCTGAGTCTCTCGTTGTTTCCCTGAGCTGTGAAGAAATCC 420
Db 949 CTTCGAGTGGGTTGCTGAGTCTCTCGTTGTTTCCCTGAGCTGTGAAGAAATCC 1008
QY 421 ACTTATGCTTCACTGATTTCTGAGCTTCTGCTTATCTGTTGCCAATTTGGCCAC 480
Db 1009 ACTTATGCTTCACTGATTTCTGAGCTTCTGCTTATCTGTTGCCAATTTGGCCAC 1068
QY 481 CGAATTTCTCCCAATCTTATCTGGCTGCGAGCGAGATGCTCTCAAGAGAGCTGATA 540
Db 1069 CGAATTTCTCCCAATCTTATCTGGCTGCGAGCGAGATGCTCTCAAGAGAGCTGATA 1128
QY 541 CAGCAGAAATGGATGGTATGTTAAATGCGACGATATACCTGTCCAAAGCTGACTTT 600
Db 1129 CAGCAGAAATGGATGGTATGTTAAATGCGACGATATACCTGTGTCCAAAGCTGACTTT 1188
QY 601 ATCCCGAGTCTCATTTCTGCGGTGCTGCTGATGACAGCTTTTGTGAAGAAATTTTG 660
Db 1189 ATCCCGAGTCTCATTTCTGCGGTGCTGCTGATGACAGCTTTTGTGAAGAAATTTTG 1248
QY 661 CCGTGTGTGAACAAATCATAGATTTTATGAGAAAGCAAAAGCTCCCAATGAGATGTT 720
Db 1249 CCGTGTGTGAACAAATCATAGATTTTATGAGAAAGCAAAAGCTCCCAATGAGATGTT 1308
QY 721 CTAGTCACTGTTAGCTGGAGTCTCCGCTCGGACCATGCTTATGCTCATCATG 780
Db 1309 CTAGTCACTGTTAGCTGGAGTCTCCGCTCGGACCATGCTTATGCTCATCATG 1368
QY 781 AAGAGATGACATGCTTTAATGAGAGCTTACAGATTTGTGAAGAAAAAGACCTACT 840
Db 1369 AAGAGATGACATGCTTTAATGAGAGCTTACAGATTTGTGAAGAAAAAGACCTACT 1428
QY 841 AATATCCCAACTTCAATTTTCTGGGCGCACTCGGATATGAGAGAAAGATTAAGAC 900
Db 1429 AATATCCCAACTTCAATTTTCTGGGCGCACTCGGATATGAGAGAAAGATTAAGAC 1488
QY 901 CAGACTGAGCATCAAGGCGCAAGAGCAACTCAAGCTGCTGACCTGAGAGAGCCAAAT 960

```
Dh 1489 CAGACTGAGCATGAGGCGCAAGCAAACTCAAGCTGCTGCACTTGAGAAAGCCAAAT 1548
Qy 961 GAACCTGTCCTGCTGCTCTGAGAGGTGAGACAGAAAGCGAGAGCGCCCTCACTCCACCC 1020
Db 1549 GAACCTGTCCTGCTGCTCTGAGAGGTGAGACAGAAAGCGAGAGCGCCCTCACTCCACCC 1608
Qy 1021 TGTGCCGACTCTGCTTACCTCTGAGAGGACAGACAAAGGCGGTGCACTCCGCGACGGT 1080
Db 1609 TGTGCCGACTCTGCTTACCTCTGAGAGGACAGACAAAGGCGGTGCACTCCGCGACGGT 1668
Qy 1081 CCCAGCGTCCGCGAGAGGCGGTGCTGCTGCTGAGAGACAGCGCGCTGTGACAGGCGCTC 1140
Db 1669 CCCAGCGTCCGCGAGAGGCGGTGCTGCTGCTGAGAGACAGCGCGCTGTGACAGGCGCTC 1728
Qy 1141 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCGTTCCTTC 1200
Db 1729 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCGTTCCTTC 1788
Qy 1201 TCTCTGATATCAAAATCAATTCATATTCAGCCAGACATGCGATCTTACATGCTTC 1260
Db 1789 TCTCTGATATCAAAATCAATTCATATTCAGCCAGACATGCGATCTTACATGCTTC 1848
Qy 1261 TCCCTCATCAGAAAGTCTTGGAAATCTCAAACTTCCACTTACTCTGATGAGACCAAC 1320
Db 1849 TCCCTCATCAGAAAGTCTTGGAAATCTCAAACTTCCACTTACTCTGATGAGACCAAC 1908
Qy 1321 AAGCTATGCCAGTTCTCCCTGTTTCAGAACTATCCGAGACAGCTCCGAAACCAAGTCT 1380
Db 1909 AAGCTATGCCAGTTCTCCCTGTTTCAGAACTATCCGAGACAGCTCCGAAACCAAGTCT 1968
Qy 1381 GATTAAGAGAGAACCCAGCATCCCCAGAAAGCTGACAGCCGCGCTTCAGACAGCCAG 1440
Db 1969 GATTAAGAGAGAACCCAGCATCCCCAGAAAGCTGACAGCCGCGCTTCAGACAGCCAG 2028
Qy 1441 AGCAAGCATTTGCTTCGGTTCAGAAACAGCAGAGAGGACCGCCAGAGGTCCTTTTA 1500
Db 2029 AGCAAGCATTTGCTTCGGTTCAGAAACAGCAGAGAGGACCGCCAGAGGTCCTTTTA 2088
Qy 1501 TCTCCACTGATCGAAATGGAGACGTTGAGAGCAATTAACACACAGCTTCCTTTTCGGC 1560
Db 2089 TCTCCACTGATCGAAATGGAGACGTTGAGAGCAATTAACACACAGCTTCCTTTTCGGC 2148
Qy 1561 CTTTCCACAGCAGACGACCTCAAGAGTGTGCTGGCTTGAGGCTTAAAGGCTTGCGAC 1620
Db 2149 CTTTCCACAGCAGACGACCTCAAGAGTGTGCTGGCTTGAGGCTTAAAGGCTTGCGAC 2208
Qy 1621 TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGCAGCTGGATTTGGC 1680
Db 2209 TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGCAGCTGGATTTGGC 2268
Qy 1681 ACAGAGTCTCTCACTTCTACTCTGCTCAGCAGCTTACGAGAGGACGTCAGATTACTT 1740
Db 2269 ACAGAGTCTCTCACTTCTACTCTGCTCAGCAGCTTACGAGAGGACGTCAGATTACTT 2328
Qy 1741 GCTTACAGTGCAGCAGCTGCTGCCACTTGGCGGAGACCAAGCTATTTCTGTGCGAGCGG 1800
Db 2329 GCTTACAGTGCAGCAGCTGCTGCCACTTGGCGGAGACCAAGCTATTTCTGTGCGAGCGG 2388
Qy 1801 CAGAAAGCAAGTACAGAGCTGACTCGGCGGAGAGCTGGACATGAAGAGAGCCCTTTGAA 1860
Db 2389 CAGAAAGCAAGTACAGAGCTGACTCGGCGGAGAGCTGGACATGAAGAGAGCCCTTTGAA 2448
Qy 1861 AAGCAGTTTAAAGCAGAGCTGCGCAATTTGAGAGAGCAGCATCATGTACAGAGAC 1920
Db 2449 AAGCAGTTTAAAGCAGAGCTGCGCAATTTGAGAGAGCAGCATCATGTACAGAGAC 2508
Qy 1921 AGGTACCGGAAAGAGCTGGGAAAGTGGGACGTACGTAGCTTTTGGGCGAGATGGAA 1980
Db 2509 AGGTACCGGAAAGAGCTGGGAAAGTGGGACGTACGTAGCTTTTGGGCGAGATGGAA 2568
Qy 1981 ATCATTGAGGTCTCC 1995
Db 2569 ATCATTGAGGTCTCC 2583
```

```
RESULT 8
AF506796
LOCUS
DEFINITION Homo sapiens dual specificity phosphatase 16 (DUSP16) mRNA.
ACCESSION AF506796
VERSION AF506796.1 GI:25573087
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3566)
Hoonmaert, I., Marynen, P., Goris, J., Sciote, R. and Baens, M.
MAPK phosphatase DUSP16/MKP-7, a candidate tumor suppressor for
chromosome region 12p12-13, reduces BCR-ABL-induced transformation
Oncogene 22 (49), 7728-7736 (2003)
14586399
JOURNAL PUBMED
2 (bases 1 to 3566)
Hoonmaert, I., Marynen, P. and Baens, M.
Direct Submision
Submitted (26-APR-2002) Department for Human Genetics-Flanders
Interuniversity Institute for Biotechnology (VIB), Katholieke
Universiteit Leuven, Herestraat 49, Leuven B-3000, Belgium
FEATURES
source
1..3566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13-p12"
/tissue_type="bone marrow"
1..3566
/gene="DUSP16"
/feature="synonym: MKP7"
633..2630
/gene="DUSP16"
/EC_number="3.1.3.48"
/EC_number="3.1.3.16"
/codon_start=1
/product="dual specificity phosphatase 16"
/protein_id="AAN75120.1"
/db_xref="GI:25573087"
/translation="MAHMEIGQIVTERLVALLSSGTRKVLIDSRPFVEYNTSHILE
AININCSKLMKRRLOQDVLTTELQHSARKKVDIDGQKVVYDQSDVAISLSDC
FLTVLLGKLEKSFNSVHLAAGFAEFSPGLCEGKSTLPTCISQPLPVANIGPT
RLPLNLYLGCQRDVNLKELMOONGIGVLANSNICPKPDIPESHFLRPVNDSECEK
ILPWLDSKVDLEKAKASNGCVLVHCLAGISRSATLIALYIMKMDMSLDAARFAVE
KRPITSPNPNFLGOLLDYERKKIKNOTGASGKSKLKLHLLEKPNRPVPAVSEGOXSB
TPLSPRCDAATSEMAAGORPVHPASVBPVSVQSLDESLVVALGLHLSARLSD
SNKLRKRSILDIKSYISASMAASLHFRSSSEDLAEYKRSYTTLDGINKLCQFSPVB
LSBQTPETSPKREKASLIPKLTQIARPSDSQSKRLHSVYTSSTGTAQSLSLPLRSGS
VEDNYHTFLFLGSLTSQHLTKSAGLGLKWHMSDILAPQISTPSLTSTSWFATESSHF
YSASAIYGGASYSAYSCSLPTCGDQVYSVRRKQKSDRADSRWSHESRSPFEKQPK
RRSCMEFGEISIMENRREHIGKVGSQSPSGMEIIVS"
ORIGIN
Query Match 99.8%; Score 1991.8; DB 9; Length 3566;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGCCCATGAGATGATGAACTCAATTTGTACTGAGAGGTGGTGGCTGCTGCGTGA 60
Db 633 ATGCCCATGAGATGATGAACTCAATTTGTACTGAGAGGTGGTGGCTGCTGCGTGA 692
Qy 61 AGTGAACGGAAGAAAGTGTCTTAATTGATAGCCGCGCATTTTGTGAATCAATATACATCC 120
Db 693 AGTGAACGGAAGAAAGTGTCTCTAATTGATAGCCGCGCATTTTGTGAATCAATATACATCC 752
Qy 121 CACATTTTGAAGCAATTAATATCACTGCTCCAGCTTATGAAGCAAGTTGCAACAG 180
```

Db 753 CACATTTGGAAACCATTAATATCACTGCTCCAGCTTATGAAAGCAAGATTGCAACAG 812
Qy 181 GAAAGATGTTATTAACAGAGCTCATCCAGCATTTCCAGGAAACATAAGTTGACATTGAT 240
Db 813 GAAAGATGTTATTAACAGAGCTCATCCAGCATTTCCAGGAAACATAAGTTGACATTGAT 872
Qy 241 TGCAGTCAGAAAGTTGATTTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 300
Db 873 TGCAGTCAGAAAGTTGATTTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 932
Qy 301 GACTGTTTCTCACTGATCTTGGGTAACTGGAAGAGCTTCACTCTGTTCACTG 360
Db 933 GACTGTTTCTCACTGATCTTGGGTAACTGGAAGAGCTTCACTCTGTTCACTG 992
Qy 361 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
Db 993 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 1052
Qy 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTGCTTACCTGTGCAATTTGGGCCAAC 480
Db 1053 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTGCTTACCTGTGCAATTTGGGCCAAC 1112
Qy 481 CGAATTTCTTCCAACTCTTATCTTGGCTGCGCAGCAGATGCTCTCAACAAGAGCTGATA 540
Db 1113 CGAATTTCTTCCAACTCTTATCTTGGCTGCGCAGCAGATGCTCTCTCAACAAGAGCTGATA 1172
Qy 541 CAGCAGAAATGGGATTTGTTATTTGTTAAATGCGCAGCTATACCTGTCCAAAGCTGACTTT 600
Db 1173 CAGCAGAAATGGGATTTGTTATTTGTTAAATGCGCAGCTATACCTGTCCAAAGCTGACTTT 1232
Qy 601 ATCCCGAGCTCATATTTCTGCGGTGCTCTGTAATGACAGCTTTTGTGAAAAATTTTG 660
Db 1233 ATCCCGAGCTCATATTTCTGCGGTGCTCTGTAATGACAGCTTTTGTGAAAAATTTTG 1292
Qy 661 CCGTGGTGGACAAATCAGTAGATTTCTTGAAGAAACAAAGCCTTCAATGATGTT 720
Db 1293 CCGTGGTGGACAAATCAGTAGATTTCTTGAAGAAACAAAGCCTTCAATGATGTT 1352
Qy 721 CTAGTGCACTGTTTAACTGCGGATCTCCCGCTCCGCCACATGCTATGCTTACATCATG 780
Db 1353 CTAGTGCACTGTTTAACTGCGGATCTCCCGCTCCGCCACATGCTATGCTTACATCATG 1412
Qy 781 AAGAGATGAGCAATGCTTATGATGAAAGCTTACAGATTTTGAAGAAAGAAAGCCTACT 840
Db 1413 AAGAGATGAGCAATGCTTATGATGAAAGCTTACAGATTTTGAAGAAAGAAAGCCTACT 1472
Qy 841 ATATCTTCAAACTTCAATTTTCTGGGCAACTCTGTGACTATGAAAGAAATTAAGAAC 900
Db 1473 ATATCTTCAAACTTCAATTTTCTGGGCAACTCTGTGACTATGAAAGAAATTAAGAAC 1532
Qy 901 CAGACTGAGAGATCAAGGCGCAAAAGCAAACTCAAGTGTGCACTGTGAAAGCCAAAT 960
Db 1533 CAGACTGAGAGATCAAGGCGCAAAAGCAAACTCAAGTGTGCACTGTGAAAGCCAAAT 1592
Qy 961 GAACTGTCCCTGCTGCTCAGAGGGTGAAGAGAAAGGAGGAGGCGCCCTCAGTCCACC 1020
Db 1593 GAACTGTCCCTGCTGCTCAGAGGGTGAAGAGAAAGGAGGAGGCGCCCTCAGTCCACC 1652
Qy 1021 TGTGCGCACTCTGCTACCTCAGAGGAGCAGAGCAAAAGGCGCTGTATCCCGCAGCGTG 1080
Db 1653 TGTGCGCACTCTGCTACCTCAGAGGAGCAGAGCAAAAGGCGCTGTATCCCGCAGCGTG 1712
Qy 1081 CCGAAGGTGCGCGGCTGCGAGCGCTGCTGTTAGAGAGCAGCGCGCTGTATCAGGGGCTC 1140
Db 1713 CCGAAGGTGCGCGGCTGCGAGCGCTGCTGTTAGAGAGCAGCGCGCTGTATCAGGGGCTC 1772
Qy 1141 AGTGGGTGACCTGTCCGAGAGAGGCTGGAAGAGCAATAGCTCAAGCTTCTTTC 1200
Db 1773 AGTGGGTGACCTGTCCGAGAGAGGCTGGAAGAGCAATAGCTCAAGCTTCTTTC 1832
Qy 1201 TCTCTGATATCAATCAGTTTATATATTCAGCAGCATGCGCAGATCTTATACATGCTTC 1260
Db 1833 TCTCTGATATCAATCAGTTTATATATTCAGCAGCATGCGCAGATCTTATACATGCTTC 1892

Qy 1261 TCTCTATCAGAAAGATGCTTGGAAATCTACAAACCTTCCAGCTACTGTGATGGGACCAAC 1320
Db 1893 TCTCTATCAGAAAGATGCTTGGAAATCTACAAACCTTCCAGCTACTGTGATGGGACCAAC 1952
Qy 1321 AAGCTATGCGCAGTTCTCCCTGTTTGAAGATATCGAGAGAGATCCCGAAACAGTCTT 1380
Db 1953 AAGCTATGCGCAGTTCTCCCTGTTTGAAGATATCGAGAGAGATCCCGAAACAGTCTT 2012
Qy 1381 GATAGAGAGAGAGCCAGCATCCCAAGAAAGCTGCAACCGCCAGGCTTTCAGACAGCCAG 1440
Db 2013 GATAGAGAGAGAGCCAGCATCCCAAGAAAGCTGCAACCGCCAGGCTTTCAGACAGCCAG 2072
Qy 1441 AGCAAGCGATTCGATTCGCTCAGAACACAGAGAGTGGACCGCCAGAGGTCCTTTTA 1500
Db 2073 AGCAAGCGATTCGATTCGCTCAGAACACAGAGAGTGGACCGCCAGAGGTCCTTTTA 2132
Qy 1501 TCTCCACTGCAATGAGTGGAGAGCTGAGAGCAATTAACAACAAGCTTCTTTTCGAG 1560
Db 2133 TCTCCACTGCAATGAGTGGAGAGCTGAGAGCAATTAACAACAAGCTTCTTTTCGAG 2192
Qy 1561 CTTTCCACAGCCAGCAGACACTCAGAAAGTCTGTGGCTGGCTTAAAGGCTGGCAC 1620
Db 2193 CTTTCCACAGCCAGCAGACACTCAGAAAGTCTGTGGCTGGCTTAAAGGCTGGCAC 2252
Qy 1621 TCGGATATCTGGGCCCCCGAGACCTTACCCCTTCCCTGACAGCAGAGCTGTATTTGGC 1680
Db 2253 TCGGATATCTGGGCCCCCGAGACCTTACCCCTTCCCTGACAGCAGAGCTGTATTTGGC 2312
Qy 1681 ACAGAGTCTCTCACTTCTACTCTGCTCAGCAGCATCTAAGAGAGAGTGCAGTACTCT 1740
Db 2313 ACAGAGTCTCTCACTTCTACTCTGCTCAGCAGCATCTAAGAGAGAGTGCAGTACTCT 2372
Qy 1741 GCTTACAGCTGACAGCAGCTGCTCCCACTTGGGAGACCAAGTCTATTTGTGCGCAGCGG 1800
Db 2373 GCTTACAGCTGACAGCAGCTGCTCCCACTTGGGAGACCAAGTCTATTTGTGCGCAGCGG 2432
Qy 1801 CAGAAGCCAAAGTACAGAGCTGACTGCGCGGCGAGCTGCGATGAAAGAGCGCCCTTGA 1860
Db 2433 CAGAAGCCAAAGTACAGAGCTGACTGCGCGGCGAGCTGCGATGAAAGAGCGCCCTTGA 2492
Qy 1861 AAGCAGTTTAAACGAGAGAGCTGCAATGGAATTTGAGAGAGCATATGTCAGAGAAC 1920
Db 2493 AAGCAGTTTAAACGAGAGAGCTGCAATGGAATTTGAGAGAGCATATGTCAGAGAAC 2552
Qy 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTACTTTTGGGAGAGATGGA 1980
Db 2553 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTACTTTTGGGAGAGATGGA 2612
Qy 1981 ATCATTGAGGTCTCC 1995
Db 2613 ATCATTGAGGTCTCC 2627

RESULT 9
AX374994
LOCUS AX374994 3766 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 17 from Patent WO0210363.
ACCESSION AX374994
VERSION AX374994.1 GI:19169826
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
1 Tang,Y.T., Elliott,V.S., Ramkumar,J., Yao,M.G., Burford,N.,
Wang,Y.B., Stewart,E.A., Gandhi,A.R., Peterson,C., Lee,E.A.,
Hafalia,A.J., Lu,D.A., Tribouley,C.M., Griffin,J.A., Baughn,M.R.,
Yue,H., Warren,B.A., Nguyen,D.B. and Walla,N.K.
Protein phosphatase
Patent: WO 0210363-A 17 07-FEB-2002;
Incyte Genomics, Inc. (US)

FEATURES

Location/Qualifiers
1..3766
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7480570CB1"

Query Match 99.8%; Score 1991.8; DB 6; Length 3766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGSCCATGAGATGATTGGAAGTCAATTTGTTACTGAGAGGTGAGTCTGCTGAGAA 60
DB 538 ATGSCCATGAGATGATTGGAAGTCAATTTGTTACTGAGAGGTGAGTCTGCTGAGAA 597
QY 61 AGTGAACGGAAAGAGTGTCTAATTGATTAAGCCCGCCATTGTTGAAATACATATACCC 120
DB 598 AGTGAACGGAAAGAGTGTCTAATTGATTAAGCCCGCCATTGTTGAAATACATATACCC 657
QY 121 CACATTTTGAAGCCCTTAATATCACTGCTCCAACTTATGAAAGCGAAGGTGCAACAG 180
DB 658 CACATTTTGAAGCCCTTAATATCACTGCTCCAACTTATGAAAGCGAAGGTGCAACAG 717
QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGCAATTCAGGAAACATAAGTTGACATTTGAT 240
DB 718 GACAAAGTGTATTAACAGAGCTCATCCAGCAATTCAGGAAACATAAGTTGACATTTGAT 777
QY 241 TGCAGTCGAAGAGTGTATGATTACATCAAGAGTCCCAAGATGTTGCTCTCTCTCTCA 300
DB 778 TGCAGTCGAAGAGTGTATGATTACATCAAGAGTCCCAAGATGTTGCTCTCTCTCTCA 837
QY 301 GACTGTTTCTCACTGACTCTCTGAGTAACTGGAAGAGCTTCAACTCTGTTCACTG 360
DB 838 GACTGTTTCTCACTGACTCTCTGAGTAACTGGAAGAGCTTCAACTCTGTTCACTG 897
QY 361 CTGAGAGTGGGTTTCTGAGTCTCTGTTGTTTCTGAGCTCTGAGAGAGAAATCC 420
DB 898 CTGAGAGTGGGTTTCTGAGTCTCTGTTGTTTCTGAGCTCTGAGAGAGAAATCC 957
QY 421 ACTGATGCTCACTGACTCTCTGAGTAACTGGAAGAGCTTCAACTCTGTTCACTG 480
DB 958 ACTGATGCTCACTGACTCTCTGAGTAACTGGAAGAGCTTCAACTCTGTTCACTG 1017
QY 481 CGAATCTTCCCAATCTTTATCTGAGCTCCAGAGAGTGTCTCAACAAAGAGCTGATA 540
DB 1018 CGAATCTTCCCAATCTTTATCTGAGCTCCAGAGAGTGTCTCAACAAAGAGCTGATA 1077
QY 541 CAGCAGAAATGGAGTTGATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 600
DB 1078 CAGCAGAAATGGAGTTGATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 1137
QY 601 ATCCCGAGTCAATTTCTGCTGCTGCTGATGATGACAGCTTTGTGAGAAATTTTG 660
DB 1138 ATCCCGAGTCAATTTCTGCTGCTGCTGATGATGACAGCTTTGTGAGAAATTTTG 1197
QY 661 CCGTGTGTTGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCAATGATGTGT 720
DB 1198 CCGTGTGTTGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCAATGATGTGT 1257
QY 721 CTAGTGCATGTTTACTGAGATCTCCGCTCCGCAACATCCCTATGCTTACATCAGT 780
DB 1258 CTAGTGCATGTTTACTGAGATCTCCGCTCCGCAACATCCCTATGCTTACATCAGT 1317
QY 781 AAGAGATGACATGTCTTTAGATGAGCTTAAGATTGTTGAAAGAAAAAGACCTACT 840
DB 1318 AAGAGATGACATGTCTTTAGATGAGCTTAAGATTGTTGAAAGAAAAAGACCTACT 1377
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAAGAC 900
DB 1378 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAAGAC 1437
QY 901 CAGACTGAGCAATCAGGGCCAAAGAGCAAACTCAAGCTGCTCACTTGAAGCCAAAT 960

DB 1438 CAGACTGAGCAATCAGGGCCAAAGAGCAAACTCAAGCTGCTCACTTGAAGCCAAAT 1497
QY 961 GAACTGTGCTCCGTGCTGCTCAAGAGGTGAGCAAGAAAGAGAGAGCCCTCACTGCAACC 1020
DB 1498 GAACTGTGCTCCGTGCTGCTCAAGAGGTGAGCAAGAAAGAGAGAGCCCTCACTGCAACC 1557
QY 1021 TGTGCCAATCTGCTTACTCCTGAGAGCAGAGCAAAAGGCCCTGATATCCCGACAGCTG 1080
DB 1558 TGTGCCAATCTGCTTACTCCTGAGAGCAGAGCAAAAGGCCCTGATATCCCGACAGCTG 1617
QY 1081 CCCAGCTGCTCCAGAGGTGAGAGCCCTGTTAGAGAGAGCCCTGATATCCCGACAGCTG 1140
DB 1618 CCCAGCTGCTCCAGAGGTGAGAGCCCTGTTAGAGAGAGCCCTGATATCCCGACAGCTG 1677
QY 1141 AGTGGCTGCACTGCTGCTGAGAGCAGAGTGAAGAGCAAGATTAAGTCAAGGCTTCTTC 1200
DB 1678 AGTGGCTGCACTGCTGCTGAGAGCAGAGTGAAGAGCAAGATTAAGTCAAGGCTTCTTC 1737
QY 1201 TCTCTGATATCAAAATGATTTCATATTCAGCCAGATGAGAGATCTTACATGCTTC 1260
DB 1738 TCTCTGATATCAAAATGATTTCATATTCAGCCAGATGAGAGATCTTACATGCTTC 1797
QY 1261 TCTCTCATCAGAAATGCTTTGAAATCTCAAAACCTTCACATCTGAGATGAGAGCAAC 1320
DB 1798 TCTCTCATCAGAAATGCTTTGAAATCTCAAAACCTTCACATCTGAGATGAGAGCAAC 1857
QY 1321 AAGCTATGCAAGTCTCCCTGTTGAGAACTATGAGAGCAAGTCTCCGAAACCAAGTCT 1380
DB 1858 AAGCTATGCAAGTCTCCCTGTTGAGAACTATGAGAGCAAGTCTCCGAAACCAAGTCT 1917
QY 1381 GATTAAGAGAAAGCCAGATATCCCAAGAGCTGAGACCCGACGCTTCAAGACAGAG 1440
DB 1918 GATTAAGAGAAAGCCAGATATCCCAAGAGCTGAGACCCGACGCTTCAAGACAGAG 1977
QY 1441 AGCAAGCAATTCATTTGATGAGAAACCAAGAGCTGAGACCCGACGCTTCAAGACAGAG 1500
DB 1978 AGCAAGCAATTCATTTGATGAGAAACCAAGAGCTGAGACCCGACGCTTCAAGACAGAG 2037
QY 1501 TCTCACTGATTCGAAGTGGAGGCTGAGAGCAATTAACACAGCTTCTTTTGCGG 1560
DB 2038 TCTCACTGATTCGAAGTGGAGGCTGAGAGCAATTAACACAGCTTCTTTTGCGG 2097
QY 1561 CTTTTCAACAGCCAGAGCACTTCAAGAGTCTGCTGAGCTTGAAGGCTGAGAC 1620
DB 2098 CTTTTCAACAGCCAGAGCACTTCAAGAGTCTGCTGAGCTTGAAGGCTGAGAC 2157
QY 1621 TCGATATCTTTGGCCCCCAGACCTTCAACCTTCCCTGACAGCAGCTGATTTTGGC 1680
DB 2158 TCGATATCTTTGGCCCCCAGACCTTCAACCTTCCCTGACAGCAGCTGATTTTGGC 2217
QY 1681 ACAAGTCTCACTGATTCATCTGCTGAGCACTTCAAGAGGAGCTGCAAGTTACTCT 1740
DB 2218 ACAAGTCTCACTGATTCATCTGCTGAGCACTTCAAGAGGAGCTGCAAGTTACTCT 2277
QY 1741 GCTTACAGCTGAGAGCACTTCAAGAGGAGCTTCAAGAGGAGCTTCAAGAGGAGCTG 1800
DB 2278 GCTTACAGCTGAGAGCACTTCAAGAGGAGCTTCAAGAGGAGCTTCAAGAGGAGCTG 2337
QY 1801 CAGAAACCAAGTGAAGAGCTGATCTGCGGCGGAGAGCTGAGCAATGAGAGAGCCCTTTGAA 1860
DB 2338 CAGAAACCAAGTGAAGAGCTGATCTGCGGCGGAGAGCTGAGCAATGAGAGAGCCCTTTGAA 2397
QY 1861 AAGCAATTTAAACGCAAGCTGCCAAATTTGAGAGAGCAATGATGAGAGAAC 1920
DB 2398 AAGCAATTTAAACGCAAGCTGCCAAATTTGAGAGAGCAATGATGAGAGAAC 2457
QY 1921 AAGTCAAGGAGAGAGCTGGGAGAGTGGAGCACTGACTTCTTTTCTGGAGAGATGAA 1980
DB 2458 AAGTCAAGGAGAGAGCTGGGAGAGTGGAGCACTGACTTCTTTTCTGGAGAGATGAA 2517
QY 1981 ATCATTTGAGTCTCC 1995

Db 2518 ATCATGAGCTCC 2532

RESULT 10
BD171157 4790 bp DNA linear PAT 17-JAN-2003
LOCUS Novel gene and protein encoded thereby.
DEFINITION BD171157
ACCESSION BD171157
VERSION WO 02052005-A/13.
KEYWORDS WO 02052005-A/13.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 4790)
AUTHORS Ohara, O., Nagase, T. and Nakajima, D.
TITLE Novel gene and protein encoded thereby
JOURNAL Patent: WO 02052005-A 13 04-JUL-2002;
KAZUSA DNA RESEARCH INSTITUTE, OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE
NAKAJIMA

COMMENT OS Homo sapiens (human)
FN WO 02052005-A/13
PD 04-JUL-2002
PF 20-DEC-2001 WO 2001JP011217
PR 22-DEC-2000 JP 00P 389742
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/12, C07K14/47
CC Novel gene and protein encoded thereby
FT Key Location/Qualifiers
CDS (184).. (2178).

FEATURES
source 1. 4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.8%; Score 1991.8; DB 6; Length 4790;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAACCAATTGTTACTAGAGAGTGTGCTGCTGGAA 60
Db 184 ATGGCCCATGAGATGATGGAACCAATTGTTACTAGAGAGTGTGCTGCTGGAA 243
QY 61 AGTGAACGGAAGAGTGTGCTGATGATGAGCCGCAATTTGTGAATCAATACATCC 120
Db 244 AGTGAACGGAAGAGTGTGCTGATGATGAGCCGCAATTTGTGAATCAATACATCC 303
QY 121 CACATTTTGAAGCATTAATATCAACTGCTCCAAAGTTATGAAGCAAGTTGCAACAG 180
Db 304 CACATTTTGAAGCATTAATATCAACTGCTCCAAAGTTATGAAGCAAGTTGCAACAG 363
QY 181 GACAAAGTGTAAATTAACAGAGCTATCCAGCATTCAGCAAGCAATAGTTGACATTGAT 240
Db 364 GACAAAGTGTAAATTAACAGAGCTATCCAGCATTCAGCAAGCAATAGTTGACATTGAT 423
QY 241 TGCAGTGAAGGTGTGATTTACGATCAAGCTCCCAAGATGTGCTCTCTTCA 300
Db 424 TGCAGTGAAGGTGTGATTTACGATCAAGCTCCCAAGATGTGCTCTCTTCA 483
QY 301 GACGTTTTCTCACTGATCTTGGGTTAACTGGAAGAGCTTCAACTGTTACCTG 360
Db 484 GACGTTTTCTCACTGATCTTGGGTTAACTGGAAGAGCTTCAACTGTTACCTG 543
QY 361 CTTCAGAGTGGTGTGATGATCTCTGTTGTTCCCTGGCTCTGTGAAGAAATTC 420
Db 544 CTTCAGAGTGGTGTGATGATCTCTGTTGTTCCCTGGCTCTGTGAAGAAATTC 603
QY 421 ACTCTAGTCCCTACCTGATCTTCTGAGCTTGTACCTGTGCAACATTTGGGCAACC 480
Db 604 ACTCTAGTCCCTACCTGATCTTCTGAGCTTGTACCTGTGCAACATTTGGGCAACC 663

QY 481 CGAATCTTCCCAATCTTATCTTGGCTGCAGAGATGTCCTCAACAGAGCTGATA 540
Db 664 CGAATCTTCCCAATCTTATCTTGGCTGCAGAGATGTCCTCAACAGAGCTGATA 723
QY 541 CAGCAGAAATGGATGGTTATGTTAAATGCCAGCTAATCTGTCCAAAGCTGACTTT 600
Db 724 CAGCAGAAATGGATGGTTATGTTAAATGCCAGCTAATCTGTCCAAAGCTGACTTT 783
QY 601 ATCCCGAGCTCATTTTCCGCTGTGCTGTGTAATGACAGCTTTGTGAAATTTTG 660
Db 784 ATCCCGAGCTCATTTTCCGCTGTGCTGTGTAATGACAGCTTTGTGAAATTTTG 843
QY 661 CCGTGGTGAACAATCAGTATGATTCATTTGAGAAAGCAAAAGCTCCATGATGTT 720
Db 844 CCGTGGTGAACAATCAGTATGATTCATTTGAGAAAGCAAAAGCTCCATGATGTT 903
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCAACATGCTATGCTATCATCAG 780
Db 904 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCAACATGCTATGCTATCATCAG 963
QY 781 AAGAGATGACATGCTTTAATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 840
Db 964 AAGAGATGACATGCTTTAATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 1023
QY 841 ATATCTCAAACCTCAATTTTCTGGGCAACCTGCTGATGATGAGAAAGATTAAGAAC 900
Db 1024 ATATCTCAAACCTCAATTTTCTGGGCAACCTGCTGATGATGAGAAAGATTAAGAAC 1083
QY 901 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGTGCACTGAGAAAGCAAAAT 960
Db 1084 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGTGCACTGAGAAAGCAAAAT 1143
QY 961 GAACTGTCTCTGCTGTCTCAGAGGTTGACAGAAAGAGAGAGCCCTCACTCCACC 1020
Db 1144 GAACTGTCTCTGCTGTCTCAGAGGTTGACAGAAAGAGAGAGCCCTCACTCCACC 1203
QY 1021 TGTGCGGACCTGCTCTCACTCAGAGGAGAGCAAAAGGCGGTGATCCGCGAGGCTG 1080
Db 1204 TGTGCGGACCTGCTCTCACTCAGAGGAGAGCAAAAGGCGGTGATCCGCGAGGCTG 1263
QY 1081 CCCAGGCTGCCAGCTGACAGCCGTGCTTTAGAGAGACAGCCGCTGTACAGGCGCTC 1140
Db 1264 CCCAGGCTGCCAGCTGACAGCCGTGCTTTAGAGAGACAGCCGCTGTACAGGCGCTC 1323
QY 1141 AGTGGCTGACCTGTCCGAGACAGCTGTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1324 AGTGGCTGACCTGTCCGAGACAGCTGTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1383
QY 1201 TCTCTGATATCAAAATCAGTTTCAATTCAGCCAGATGGCAGATCTTACATGGCTTC 1260
Db 1384 TCTCTGATATCAAAATCAGTTTCAATTCAGCCAGATGGCAGATCTTACATGGCTTC 1443
QY 1261 TCTCTATCAGAAAGATCTTGGAAATCTCAAACTTCCATCTATCTGTGATGGAAAC 1320
Db 1444 TCTCTATCAGAAAGATCTTGGAAATCTCAAACTTCCATCTATCTGTGATGGAAAC 1503
QY 1321 AAGCTATGCAAGTCTTCCCTGTTCAGAACTATGAGAGACACTCCGAAACAGACTCT 1380
Db 1504 AAGCTATGCAAGTCTTCCCTGTTCAGAACTATGAGAGACACTCCGAAACAGACTCT 1563
QY 1381 GATTAAGAGAGAGCAGCATCCCAAGAGCTGCAACCGCCAGGCTTCAGACAGCCAG 1440
Db 1564 GATTAAGAGAGAGCAGCATCCCAAGAGCTGCAACCGCCAGGCTTCAGACAGCCAG 1623
QY 1441 AGCAAGCAGATTCGATTCGATCAAGAACAGCAGAGTGGACCGCCAGAGGTCCTTTTA 1500
Db 1624 AGCAAGCAGATTCGATTCGATCAAGAACAGCAGAGTGGACCGCCAGAGGTCCTTTTA 1683
QY 1501 TCTCCACTGATCAGAGTGGAGAGCTGTGAGAGCAATTAACACACAGCTTCTTTTGGC 1560
Db 1684 TCTCCACTGATCAGAGTGGAGAGCTGTGAGAGCAATTAACACACAGCTTCTTTTGGC 1743
QY 1561 CTTTCACACAGCAGACGACCTCAGAAAGTCTGAGGCTTGAAGGCTGGCAC 1620

Db 1744 CTTTCACGACGACGACGACCTCAAGAGTCTGCGCTGGGCTTTAAAGGCTGGCAC 1803
QY 1621 TCGGAATATCTTGGCCCCCGCCAGACCTTACCCCTTCCCTGACACAGCTGGTATTTTGGC 1680
Db 1804 TCGGAATATCTTGGCCCCCGCCAGACCTTACCCCTTCCCTGACACAGCTGGTATTTTGGC 1683
QY 1681 ACAGAGTCTCTCAACCTTCTTACTCTGCTGACCATCTHCGAGGACAGTTCAGTTACTCT 1740
Db 1864 ACAGAGTCTCTCAACCTTCTTACTCTGCTGACCATCTHCGAGGACAGTTCAGTTACTCT 1923
QY 1741 GCTTACAGCTGACGACCTGCTCCCACTTTCGAGACCAAGTCTATTTCTGCGCAGGCGG 1800
Db 1924 GCTTACAGCTGACGACCTGCTCCCACTTTCGAGACCAAGTCTATTTCTGCGCAGGCGG 1983
QY 1801 CAGAACCCAGTGAACAGAGTGACTGCGGGCGGAGCTGCGCATGAAGAGCCCTTTGAA 1860
Db 1984 CAGAACCCAGTGAACAGAGTGACTGCGGGCGGAGCTGCGCATGAAGAGCCCTTTGAA 2043
QY 1861 AAGCATTTTAAACGACAGACCTGCGCAATGGAATTTTGAGAGAGCATCATGTCAAGAAC 1920
Db 2044 AAGCATTTTAAACGACAGACCTGCGCAATGGAATTTTGAGAGAGCATCATGTCAAGAAC 2103
QY 1921 AGGTCAAGGGAAGAGCTGCGGAAAGTGGGAGTCACTGCTTTTGGGCGAGCATGAA 1980
Db 2104 AGGTCAAGGGAAGAGCTGCGGAAAGTGGGAGTCACTGCTTTTGGGCGAGCATGAA 2163
QY 1981 ATCATTTAGAGTCTCC 1995
Db 2164 ATCATTTAGAGTCTCC 2178

RESULT 11
BD183422 4790 bp DNA linear PAT 17-JUN-2003
LOCUS BD183422
DEFINITION Novel genes and proteins encoded by the genes.
ACCESSION BD183422.1 GI:31875622
VERSION JP 2002345492-A/135.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE Ohara, O., Nagase, T. and Nakajima, D.
AUTHORS Novel genes and proteins encoded by the genes
TITLE Patent: JP 2002345492-A 135 03-DEC-2002;
JOURNAL KAZUSA DNA RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002345492-A/135
PD 03-DEC-2002
PF 26-FEB-2002 JP 2002049009
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,
PC A61P25/14, A61P35/00, C12N15/00, A61K37/02
CC Novel genes and proteins encoded by the genes FH
Location/Qualifiers
FT CDS (184)..(2178).
FEATURES
source 1..4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1991.8; DB 6; Length 4790;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGCCCCATGAGATGATTTGGAATCTCAATTTTACTGAGAGTTGGTGGCTCTGCGGAA 60
Db 184 ATGCCCCATGAGATGATTTGGAATCTCAATTTTACTGAGAGTTGGTGGCTCTGCGGAA 243

QY 61 AGTGAACGGAAGAAAGTCTGCTAAATTGATAGCCGGCATTGTTGGAATCAATACATCC 120
Db 244 AGTGAACGGAAGAAAGTCTGCTAAATTGATAGCCGGCATTGTTGGAATCAATACATCC 303
QY 121 CACATTTTGAAGCCATTATATCAATGCTCCCAAGCTTATGAAGCAAGTTGCAACAG 180
Db 304 CACATTTTGAAGCCATTATATCAATGCTCCCAAGCTTATGAAGCAAGTTGCAACAG 363
QY 181 GACAAAGTTTAAATTCAGAGCTCATTCAGCATTCAGCGAAATTAAGTTGACATTTAT 240
Db 364 GACAAAGTTTAAATTCAGAGCTCATTCAGCATTCAGCGAAATTAAGTTGACATTTAT 423
QY 241 TGCAGTGAAGAGTTGATGTTTATGATCAAGTCAAGTCCCAAGATTTGCTCTCTTCA 300
Db 424 TGCAGTGAAGAGTTGATGTTTATGATCAAGTCCCAAGATTTGCTCTCTTCA 483
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAATCTGGAAGAGCTTCAACTGTTCACTG 360
Db 484 GACTGTTTCTCACTGATCTTCTGGGTAATCTGGAAGAGCTTCAACTGTTCACTG 543
QY 361 CTTCAGAGTGGGTTTGTGCTGATCTCTGCTGTTTCCCTGGCTTGTGAAGAAATTC 420
Db 544 CTTCAGAGTGGGTTTGTGCTGATCTCTGCTGTTTCCCTGGCTTGTGAAGAAATTC 603
QY 421 ACTCTAGTCCCTACCTGACATTTCTGAGCTTGTCTTATCTGTTGCAACATTTGGCCAAC 480
Db 604 ACTCTAGTCCCTACCTGACATTTCTGAGCTTGTCTTATCTGTTGCAACATTTGGCCAAC 663
QY 481 CGAATTTCTCCCAATCTTATCTTGGCTGCGCAGAGATGCTTCAACAGAGCTGAAT 540
Db 664 CGAATTTCTCCCAATCTTATCTTGGCTGCGCAGAGATGCTTCAACAGAGCTGAAT 723
QY 541 CAGCAAGATGGATGTTGTTATGTTAAATGCGCATATCTGTGCAAGCTGACTT 600
Db 724 CAGCAAGATGGATGTTGTTATGTTAAATGCGCAATATCTGTGCAAGCTGACTT 783
QY 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 784 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843
QY 661 CCGTGGTGGCAATAGATGATGATTTCAATTTGAAAGCAAGCAAGCTTCAATGATGTT 720
Db 844 CCGTGGTGGCAATAGATGATGATTTCAATTTGAAAGCAAGCAAGCTTCAATGATGTT 903
QY 721 CTAGTGAATGTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 904 CTAGTGAATGTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACGATTTTGAAGAAAGAAAGCACTACT 840
Db 964 AAGAGATGACATGCTTTAGATGAAGCTTACGATTTTGAAGAAAGAAAGCACTACT 1023
QY 841 ATATCTCCCAATCTCAATTTTCTGGGCCAATCTCTGATATGAGAAAGATTTAAGAC 900
Db 1024 ATATCTCCCAATCTCAATTTTCTGGGCCAATCTCTGATATGAGAAAGATTTAAGAC 1083
QY 901 CAGATGAGATCAAGGCGCAAGCAAGCAAGCAAGCAAGCTGATGCTGGAAGCACTAAT 960
Db 1084 CAGATGAGATCAAGGCGCAAGCAAGCAAGCAAGCAAGCTGATGCTGGAAGCACTAAT 1143
QY 961 GAACCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 1144 GAACCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1203
QY 1021 TGTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1204 TGTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1263
QY 1081 CCGAGCTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1264 CCGAGCTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1323

QY 361 CTTCAGGTGGGTTTGTCTGAGTCTCTGCTGTTTCCCTGGGCTCTGTGAGGAAATCC 420
Db 544 CTTCAGGTGGGTTTGTCTGAGTCTCTGCTGTTTCCCTGGGCTCTGTGAGGAAATCC 603
QY 421 ACTCTAGTCCCTACCTGCAATTTCTGAGGCTTGCTTACCTGTGGCAACATTTGGGCAAC 480
Db 604 ACTCTAGTCCCTACCTGCAATTTCTGAGGCTTGCTTACCTGTGGCAACATTTGGGCAAC 663
QY 481 CGAATTTCTTCCCAATCTTATCTTGGGCTGCAAGAGATGTCTTCAACAAGAGCTGATG 540
Db 664 CGAATTTCTTCCCAATCTTATCTTGGGCTGCAAGAGATGTCTTCAACAAGAGCTGATG 723
QY 541 CAGCAATTTGGGATTTGGTATGTGTAAATGCAAGCTAATCTGTCGCAAGCTGATTT 600
Db 724 CAGCAATTTGGGATTTGGTATGTGTAAATGCAAGCTAATCTGTCGCAAGCTGATTT 783
QY 601 ATCCCGAGTCTCAATTTCTGCGGTGGCTGCAATGACAGCTTTTGTGAGAAATTTTG 660
Db 784 ATCCCGAGTCTCAATTTCTGCGGTGGCTGCAATGACAGCTTTTGTGAGAAATTTTG 843
QY 661 CCGGTGTGGACAAATGATGATTTCAATTTGAGAAAGCAAAAGCTCCCAATGATTTGTT 720
Db 844 CCGGTGTGGACAAATGATGATTTCAATTTGAGAAAGCAAAAGCTCCCAATGATTTGTT 903
QY 721 CTAGTCACTGTTTACCTGAGATCTCCGCTCCGCAATGCTTATGCTATGCTATCATG 780
Db 904 CTAGTCACTGTTTACCTGAGATCTCCGCTCCGCAATGCTTATGCTATGCTATCATG 963
QY 781 AAGAGATGAGACATGCTTTAGATGATGATTTCAAGATTTGAGAAAGAAAGCACTACT 840
Db 964 AAGAGATGAGACATGCTTTAGATGATGATTTCAAGATTTGAGAAAGAAAGCACTACT 1023
QY 841 ATATCTCCAACTTCAATTTCTGGGCAATCTGCTGCACTATGAGAAAGAAAGCACTACT 900
Db 1024 ATATCTCCAACTTCAATTTCTGGGCAATCTGCTGCACTATGAGAAAGAAAGCACTACT 1083
QY 901 CAGACTGGAGCATCAGGGGCAAAAGCAAACTCAAGCTGTGCACTGAGAGAGCAAAAT 960
Db 1084 CAGACTGGAGCATCAGGGGCAAAAGCAAACTCAAGCTGTGCACTGAGAGAGCAAAAT 1143
QY 961 GAACCTGTCCCTGTGTCTGAGAGGTGAGACAGAAAGGAGACCGCTCTCACTCAACC 1020
Db 1144 GAACCTGTCCCTGTGTCTGAGAGGTGAGACAGAAAGGAGACCGCTCTCACTCAACC 1203
QY 1021 TGTCCGCACTGTCTACCTCAGAGGCAAGCAAGCAAAAGGCGCTCATCCGCAAGGCTG 1080
Db 1204 TGTCCGCACTGTCTACCTCAGAGGCAAGCAAGCAAAAGGCGCTCATCCGCAAGGCTG 1263
QY 1081 CCCAGGCTGCCAGCGCTGAGCGCTGTGTTAGAGAGACAGCGGCTGTGACAGGCTTC 1140
Db 1264 CCCAGGCTGCCAGCGCTGAGCGCTGTGTTAGAGAGACAGCGGCTGTGACAGGCTTC 1323
QY 1141 AGTGGCTGCACTGTCCGCAAGGCTGAGAGCAAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1324 AGTGGCTGCACTGTCCGCAAGGCTGAGAGCAAGCAATTAAGCTCAAGCTTCTTC 1383
QY 1201 TCTCTGGAATTAATCAATCAATTTCAATTTCAAGCAAGTGGAGATCTTACATGCTTC 1260
Db 1384 TCTCTGGAATTAATCAATCAATTTCAATTTCAAGCAAGTGGAGATCTTACATGCTTC 1443
QY 1261 TCTCTGATGAGATGCTTTGGAATTAATCAAACTTCACTCTGATGGAGCAAC 1320
Db 1444 TCTCTGATGAGATGCTTTGGAATTAATCAAACTTCACTCTGATGGAGCAAC 1503
QY 1321 AAGCTATGCCAGTTTCTCCCTGTTTCAAGAACTATCGAGAGCACTCCGAAACAGTCTCT 1380
Db 1504 AAGCTATGCCAGTTTCTCCCTGTTTCAAGAACTATCGAGAGCACTCCGAAACAGTCTCT 1563
QY 1381 GATTAAGAGAGAGCAAGCTCCCAAGAGCTGCAAGCGCTTTCAGAGAGCGAC 1440
Db 1564 GATTAAGAGAGAGCAAGCTCCCAAGAGCTGCAAGCGCTTTCAGAGAGCGAC 1623

QY 1441 AGCAAGCATTTGATTTGGTCCAGAAACAGACAGATGACACCGCCAGAGGTCCCTTTTA 1500
Db 1624 AGCAAGCATTTGATTTGGTCCAGAAACAGACAGATGACACCGCCAGAGGTCCCTTTTA 1683
QY 1501 TCTCCACTGATTCGAAATGGAGAGCGTGGAGCAATTAACACACAGCTTCTTTTGGC 1560
Db 1684 TCTCCACTGATTCGAAATGGAGAGCGTGGAGCAATTAACACACAGCTTCTTTTGGC 1743
QY 1561 CTTTCACAGCAGCAGACACCTCAAGAGTCTGCTGCTGGGCTTTAAGGCTGGCAC 1620
Db 1744 CTTTCACAGCAGCAGACACCTCAAGAGTCTGCTGCTGGGCTTTAAGGCTGGCAC 1803
QY 1621 TGGATATCTTGGGCTCCGAGACCTTACCTTCCCTGACAGACAGCTGATTTTGGC 1680
Db 1804 TGGATATCTTGGGCTCCGAGACCTTACCTTCCCTGAGAGGAGGAGCTGATTTTGGC 1863
QY 1681 ACAGAGTCTCACACTTTCTACTGTGCTGAGGCACTTACAGAGGAGGAGCTGATTTCT 1740
Db 1864 ACAGAGTCTCACACTTTCTACTGTGCTGAGGCACTTACAGAGGAGGAGCTGATTTCT 1923
QY 1741 GCTTACAGCTGAGCCAGCTGCTCCACTTGGGAGACCAAGTCTATTTGTGGAGAGG 1800
Db 1924 GCTTACAGCTGAGCCAGCTGCTCCACTTGGGAGACCAAGTCTATTTGTGGAGAGG 1983
QY 1801 CAGAGGCAAGTGAAGAGCTGACTGCGGCGGAGCTGGCATGAAAGAGCCCTTTGAA 1860
Db 1984 CAGAGGCAAGTGAAGAGCTGACTGCGGCGGAGCTGGCATGAAAGAGCCCTTTGAA 2043
QY 1861 AAGCATTTAAACGAGAAAGCTGCCAAATGAAATTTGGAGAGACATCATGTGAGAAC 1920
Db 2044 AAGCATTTAAACGAGAAAGCTGCCAAATGAAATTTGGAGAGACATCATGTGAGAAC 2103
QY 1921 AGGTCAAGGAAAGCTGGGAAAGTGGGCACTGACTGCTTTTGGGCAAGTGA 1980
Db 2104 AGGTCAAGGAAAGCTGGGAAAGTGGGCACTGACTGCTTTTGGGCAAGTGA 2163
QY 1981 ATCATTTAGGCTTCC 1995
Db 2164 ATCATTTAGGCTTCC 2178

RESULT 13
AXI80875 2732 bp DNA linear PAT 06-AUG-2001
LOCUS AXI80875
DEFINITION Sequence 2 from Patent WO0146394.
ACCESSION AXI80875
VERSION AXI80875.1 GI:15132703
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
Plowman, G.D., Martinez, R., Whyte, D., Manning, G., Sudarshan, S.,
Hill, R.J., and Flanagan, P.,
Mammalian protein phosphatases
Patent: WO 0146394-A 2 28-JUN-2001;
Sugen, Inc. (US)

FEATURES
source
1..2732
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1990.2; DB 6; Length 2732;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATTTGGAACCTCAATTTGTTACTGAGAGGTTGGTGTCTGTGGA 60
Db 538 ATGGCCCATGAGATGATTTGGAACCTCAATTTGTTACTGAGAGGTTGGTGTCTGTGGA 597

QY AGTGAACGAAAAAGTGTCTTAATTGATAGCGGCACTTTGNGAATAACATCC 120
DB AGTGAACGAAAAAGTGTCTCTTAATTGATAGCGGCACTTTGNGAATAACATCC 657
QY CACATTTGGAAAGCATTATATCACTGCTCCAGGTTATGAAGGAGTTGCAACG 180
DB CACATTTGGAAAGCATTATATCACTGCTCCAGGTTATGAAGGAGTTGCAACG 717
QY 181 GACAAAGTGTATATACAGGCTATCCAGCATTCAGGAAACATTAAGTTGACATTTGAT 240
DB GACAAAGTGTATATACAGGCTATCCAGCATTCAGGAAACATTAAGTTGACATTTGAT 777
QY 241 TGCAGTCAGAAAGTGTATATGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
DB TGCAGTCAGAAAGTGTATATGATCAAGCTCCCAAGATGTTGCTCTCTCTCTCA 837
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAGAGAGCTTCACTCTGTTCACTG 360
DB GACTGTTTCTCACTGATCTTCTGGGTAACTGGAGAGAGCTTCACTCTGTTCACTG 897
QY 361 CTTCAGAGTGGTGTGCTGAGTCTCTGTTGTTTCTGCTGCTCTGTAAGAAATCC 420
DB CTTCAGAGTGGTGTGCTGAGTCTCTGTTGTTTCTGCTGCTCTGTAAGAAATCC 957
QY 421 ACTCTAGTCCCTCACTGCACTTCTGAGCTTGTACCTGTTGCAACATTTGGGCAACC 480
DB ACTCTAGTCCCTCACTGCACTTCTGAGCTTGTACCTGTTGCAACATTTGGGCAACC 1017
QY 481 CGAATTTCTTCCAACTTTATCTTGGCTGCGAGAGAGATGCTTCAACAGAGCTGATA 540
DB CGAATTTCTTCCAACTTTATCTTGGCTGCGAGAGAGATGCTTCAACAGAGCTGATA 1018
QY 541 CAGCAGAAATGGATGGTTATGTTTAAATGCGAGCTATACCTGTTCCAAAGCTGACTT 600
DB CAGCAGAAATGGATGGTTATGTTTAAATGCGAGCTATACCTGTTCCAAAGCTGACTT 1078
QY 601 ATCCCGAGTCTCAATTTCTGCGGTGCTGTAAGTGAAGAGCTTTGTGAAGAAATTTTG 660
DB ATCCCGAGTCTCAATTTCTGCGGTGCTGTAAGTGAAGAGCTTTGTGAAGAAATTTTG 1138
QY 661 CCGTGTGTGACAAATCAGTGAATTTCAATGAGAAAGCAAAAGCTTCCAAATGATGTT 720
DB CCGTGTGTGACAAATCAGTGAATTTCAATGAGAAAGCAAAAGCTTCCAAATGATGTT 1198
QY 721 CTATGCACTGTTAGTGGATCTCCGCTCGCCCACTCGCTATCGCTATCATCATG 780
DB CTATGCACTGTTAGTGGATCTCCGCTCGCCCACTCGCTATCGCTATCATCATG 1258
QY 781 AAGAGATGAGCAATGCTTAAATGAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 840
DB AAGAGATGAGCAATGCTTAAATGAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 1318
QY 841 ATATCTTCAAACTTCAATTTCTGGGCAATCTCTGAGCTATGAGAAAGATTAAGAAC 900
DB ATATCTTCAAACTTCAATTTCTGGGCAATCTCTGAGCTATGAGAAAGATTAAGAAC 1378
QY 901 CAGACTGAGCATCAGGGGCCAAAGAGCAATCAAGTGTGCACTGTGAAGAACCAAT 960
DB CAGACTGAGCATCAGGGGCCAAAGAGCAATCAAGTGTGCACTGTGAAGAACCAAT 1438
QY 961 GAACCTGTCTCTGTCTGAGAGGTGAGCAGAAAGCGAGGCGCTCACTGCAACC 1020
DB GAACCTGTCTCTGTCTGAGAGGTGAGCAGAAAGCGAGGCGCTCACTGCAACC 1498
QY 1021 TGTGCGCACTCTGCTACTCAGAGGAGCAGAGCAAAAGGCGCTGATCCCGCAGGCTG 1080
DB TGTGCGCACTCTGCTACTCAGAGGAGCAGAGCAAAAGGCGCTGATCCCGCAGGCTG 1558
QY 1081 CCCAGCTGCGCAGCGTGCCTGTTAGAGAGCAGCCGCTGTTAGCAGGCGCTC 1140
DB CCCAGCTGCGCAGCGTGCCTGTTAGAGAGCAGCCGCTGTTAGCAGGCGCTC 1618
QY 1141 AGTGGGCTGACCTGTGCGGAGCAGGCTGGAAGAGCAATTAAGCTCAAGCGTTCTTC 1200

DB 1678 AGTGGGCTGACCTGTGCGGAGCAGGCTGGAAGAGCAAGCAATTAAGCTCAAGCTTCTTC 1737
QY 1201 TCTTGTGATATCAATCACTTCTATATTAAGCCAGATGAGCATCTTACATGCTTC 1260
DB TCTTGTGATATCAATCACTTCTATATTAAGCCAGATGAGCATCTTACATGCTTC 1738
QY 1261 TCTTCAATCAAGAGCTTGGAAATCTAACAACCTTCACTGCTGATGAGGAGCAAC 1320
DB TCTTCAATCAAGAGCTTGGAAATCTAACAACCTTCACTGCTGATGAGGAGCAAC 1798
QY 1321 AAGCTATGCACTTCTCCCTGTTGAGAACTATCGAGAGAGCTCCGAAACAGTCTT 1380
DB AAGCTATGCACTTCTCCCTGTTGAGAACTATCGAGAGAGCTCCGAAACAGTCTT 1858
QY 1381 GATAGAGAGAGAGCCAGCATTCGCCAAGAAAGCTGAGACCGCCAGGCTTCAAGACCCAG 1440
DB GATAGAGAGAGAGCCAGCATTCGCCAAGAAAGCTGAGACCGCCAGGCTTCAAGACCCAG 1918
QY 1441 AGCAAGCAATTCAGTTCAGTTCAGAAACAGCAGAGTGGGACCGGCGAGAGTCTTTTA 1500
DB AGCAAGCAATTCAGTTCAGTTCAGAAACAGCAGAGTGGGACCGGCGAGAGTCTTTTA 1978
QY 1501 TCTCACTGCACTGAAAGTGGAGAGCTGAGAGCAATTAACACACAGCTTCTTTCGCG 1560
DB TCTCACTGCACTGAAAGTGGAGAGCTGAGAGCAATTAACACACAGCTTCTTTCGCG 2038
QY 1561 CTTCACACAGCAGCAGACCTCAGAAAGTGTGCTGAGGCTTGAAGGCTGAGCAC 1620
DB CTTCACACAGCAGCAGACCTCAGAAAGTGTGCTGAGGCTTGAAGGCTGAGCAC 2098
QY 1621 TGGATATCTGGGCGCCGAGACCTTACCCCTTCCCTGACGAGAGCTGATTTTGGC 1680
DB TGGATATCTGGGCGCCGAGACCTTACCCCTTCCCTGACGAGAGCTGATTTTGGC 2158
QY 1681 ACAGAGTCTCACTTCTACTCTGCTCAGCAGCATCTAAGAGGAGTGCATTACTCT 1740
DB ACAGAGTCTCACTTCTACTCTGCTCAGCAGCATCTAAGAGGAGTGCATTACTCT 2218
QY 1741 GCTTACAGCTGACAGCAGCTGCTCAGTGGAGACCAAGTCTTATCTGTGCGCAGGCG 1800
DB GCTTACAGCTGACAGCAGCTGCTCAGTGGAGACCAAGTCTTATCTGTGCGCAGGCG 2278
QY 1801 CAGAAAGCAGAGCAGAGCTGAGCTGCGGAGAGTGGAGATGAAGAGAGCCCTTGA 1860
DB CAGAAAGCAGAGCAGAGCTGAGCTGCGGAGAGTGGAGATGAAGAGAGCCCTTGA 2338
QY 1861 AAGCAGTTTAAAGCAGAGCTGCAATTAAGTGAATTTGAGAGAGCATGTCAGAGAAC 1920
DB AAGCAGTTTAAAGCAGAGCTGCAATTAAGTGAATTTGAGAGAGCATGTCAGAGAAC 2398
QY 1921 AGGTCAAGGAGAGCTGGGAGAAAGTGGGAGTCACTGATCTTTCGGGAGCAGATGAA 1980
DB AGGTCAAGGAGAGCTGGGAGAAAGTGGGAGTCACTGATCTTTCGGGAGCAGATGAA 2458
QY 1981 ATCATTAAGTCTCC 1995
DB ATCATTAAGTCTCC 2518
RESULT 14
AX405700 3104 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 115 from Patent W00222660.
DEFINITION AX405700
ACCESSION AX405700.1 GI:21438839
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,

TITLE
Xue,A.-J., Yang,Y., Wehrman,T. and Drmanac,R.T.
Novel nucleic acids and polypeptides
Patent: WO 0222660-A 115 21-MAR-2002;
JOURNAL
HYSO, INC. (US)

FEATURES

Location/Qualifiers
1..3104
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
159..2156
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD34807.1"
/db_xref="GI:21438840"
/db_xref="RENTREMBL:CAD34807"
/translation="MAHEMIGTIVTEELVLLBSGETKVLIDISPEFVEXTSHILE
AIVNCSKLMKRLIODEKVLITELIYHAKHAKVNDISQKVVVDOSQOVASISPC
FLTVLLGLEKSPNSVYLLAGFAFSCFPGLEKSTLVPTCISOPCLPVANIGPT
RLPVLXGCCORDVYNKELMOQNGIGVLANSTCPKPDTPESHLFRLVPVNDSPCK
ILPMLDKSYDFLEKASNSGCVLVHCLAGISRSATIAIAYIMKMDMLAEAVRPIKE
KAPTISVNFNFIQGLIDYEKIKNOTGASGPKSLKILHEKPEPVPAYBSGQKSE
TLPSPCADSATSEAGORPVHPASVPSVPSVPSLEDSPLQALSLHLSADRLED
SNKLRPSLDIKSVYSASMAAHLHSPSSEDALRYKSTTLIDGINKLQSPVQOE
LSEQTPETSPDEKASIPKIKOTARPSPSOSKRLHVRISSSGTAORSLSPHRSQS
VEDNYHTEFLRGLSTSOOHLKSMAGLKGHMSDILAPORTSPSLTWSWTFATESHF
YSAAIYGGASYSYSSQPLPTGDDYIVVRROKPSDRADRSRSHESPEKQFK
RRSCMERGESIMENSRREELRGVGSQSSPSGSMELIEVS"

CDS

ORIGIN

Query Match 99.8%; Score 1990.2; DB 6; Length 3104;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 ATGGCCCATGAGATGATGGAATCACTCAATTTGTACTGAGAGTGGTGGCTCTGCTGAA 60
Db ATGGCCCATGAGATGATGGAATCACTCAATTTGTACTGAGAGTGGTGGCTCTGCTGAA 60
QY 159 ATGGCCCATGAGATGATGGAATCACTCAATTTGTACTGAGAGTGGTGGCTCTGCTGAA 218
QY 61 AGTGAACGGAAAAAGTCTGCTAATGATGAGCGGCGCAATTTGGAATACAAATCATCC 120
Db 219 AGTGAACGGAAAAAGTCTGCTAATGATGAGCGGCGCAATTTGGAATACAAATCATCC 278
QY 121 CACATTTTGGAGCGCATTAATCACTGCTCCAGCTTATGAGCGAGGTGGCAAG 180
Db 279 CACATTTTGGAGCGCATTAATCACTGCTCCAGCTTATGAGCGAGGTGGCAAG 338
QY 181 GACAAAGTTAATTAACAGAGTCATCCAGCAATTCAGCGAAACATTAAGTTGCAATTGAT 240
Db 339 GACAAAGTTAATTAACAGAGTCATCCAGCAATTCAGCGAAACATTAAGTTGCAATTGAT 398
QY 241 TGCAGTCAGAGGTTGATGTTACGATCAAGAGTCCCAAGATGTTGCTCTCTCTTCA 300
Db 399 TGCAGTCAGAGGTTGATGTTACGATCAAGAGTCCCAAGATGTTGCTCTCTCTTCA 458
QY 301 GACTGTTTTCTCACTGTAATCTTCTGGGTAACTGAGGAAGAGCTTCAACTCTGTTCACTTG 360
Db 459 GACTGTTTTCTCACTGTAATCTTCTGGGTAACTGAGGAAGAGCTTCAACTCTGTTCACTTG 420
QY 361 CTTCGAGGAGGTTGCTGAGATCTCTGCTGTTTCCCTGGCCCTCTGTCGAAGAAATCC 518
Db 519 CTTCGAGGAGGTTGCTGAGATCTCTGCTGTTTCCCTGGCCCTCTGTCGAAGAAATCC 578
QY 421 ACTCTAGTCCCTCACTGCAATTTCTCAGCTTGGCTTAACTGTCGCAACATTTGGCCCAAC 480
Db 579 ACTCTAGTCCCTCACTGCAATTTCTCAGCTTGGCTTAACTGTCGCAACATTTGGCCCAAC 638
QY 481 CGAATTTCTCCCATCTTTATCTTGGCTGCGCAGCGAGATGTCCTCAACAAGAGGCTGATA 540
Db 639 CGAATTTCTCCCATCTTTATCTTGGCTGCGCAGCGAGATGTCCTCAACAAGAGGCTGATA 698
QY 541 CAGCAGATGGGATGGTATGTTAAATGCAAGCTATACCTGTCGAAAGCTGACCTT 600
Db 699 CAGCAGATGGGATGGTATGTTAAATGCAAGCTATACCTGTCGAAAGCTGACCTT 758

QY 601 ATCCCCGAGTTCATTTCTGCGTGTGCTGGAATGACAGCTTTTGAGAAATTTTG 660
Db 759 ATCCCCGAGTTCATTTCTGCGTGTGCTGGAATGACAGCTTTTGAGAAATTTTG 818
QY 661 CCGTGGTTGACAATATCATGATGATTTCTTGAAGAAAGCAAAAGCTTCAATGATGTT 720
Db 819 CCGTGGTTGACAATATCATGATGATTTCTTGAAGAAAGCAAAAGCTTCAATGATGTT 878
QY 721 CTATGCACTGTTTATGCTGGATCTCCGCTCCGCACTTCGCTATGCTATCATATG 780
Db 879 CTATGCACTGTTTATGCTGGATCTCCGCTCCGCACTTCGCTATGCTATCATATG 938
QY 781 AAGAGATGAGCAATGCTTTTGAATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 840
Db 939 AAGAGATGAGCAATGCTTTTGAATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 998
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGCAATGAGAAAGAAATTAAGAAC 900
Db 999 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGCAATGAGAAAGAAATTAAGAAC 1058
QY 901 CAGACTGAGATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGAGAGAGCAAT 960
Db 1059 CAGACTGAGATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGAGAGAGCAAT 1118
QY 961 GAACCTGCTCCTGCTGCTCAGAGAGGTGAGCAAGAAAGGAGAGAGCCCTCAATCACC 1020
Db 1119 GAACCTGCTCCTGCTGCTCAGAGAGGTGAGCAAGAAAGGAGAGAGCCCTCAATCACC 1178
QY 1021 TGGCGGACTGCTGCTCCTCAAGAGCAGAGCAAAAGCCCTGTCATCCCGCAGAGTG 1080
Db 1179 TGGCGGACTGCTGCTCCTCAAGAGCAGAGCAAAAGCCCTGTCATCCCGCAGAGTG 1238
QY 1081 CCGAGGTCGCCAGGTCAGGCTGCTGCTGCTGTAAGAGACACCCGCTGTAACAGGCTC 1140
Db 1239 CCGAGGTCGCCAGGTCAGGCTGCTGCTGCTGTAAGAGACACCCGCTGTAACAGGCTC 1298
QY 1141 AGTGGCTGCACTTCTCCGCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1299 AGTGGCTGCACTTCTCCGCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1358
QY 1201 TCTTGGATATCAATTCAGTTTCAATTTAGCCGAGCAGATGAGCAATTCATGCTTC 1260
Db 1359 TCTTGGATATCAATTCAGTTTCAATTTAGCCGAGCAGATGAGCAATTCATGCTTC 1418
QY 1261 TCTCATGAGAAAGTCTTGAATTAATCAAACTTTCATCTGATGAGAGCAAC 1320
Db 1419 TCTCATGAGAAAGTCTTGAATTAATCAAACTTTCATCTGATGAGAGCAAC 1478
QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATGAGAGAGATTCGCCAAACAGTCTCT 1380
Db 1479 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATGAGAGAGATTCGCCAAACAGTCTCT 1538
QY 1381 GATTAAGAGAGAGCCAGCATCTCCCAAGAAAGCTGACAGCCGAGCTTTCAGACAGCCAG 1440
Db 1539 GATTAAGAGAGAGCCAGCATCTCCCAAGAAAGCTGACAGCCGAGCTTTCAGACAGCCAG 1598
QY 1441 AGCAAGGATTTGATTCGATCAGAAACAGCAGAGTGGACCCGCAAGAGTCCCTTTTA 1500
Db 1599 AGCAAGGATTTGATTCGATCAGAAACAGCAGAGTGGACCCGCAAGAGTCCCTTTTA 1658
QY 1501 TCTCCAGTGCATGAGAGTGGAGGCGTGAAGCAATTAACAACAAGCTTCTTTTGGG 1560
Db 1659 TCTCCAGTGCATGAGAGTGGAGGCGTGAAGCAATTAACAACAAGCTTCTTTTGGG 1718
QY 1561 CTCTCCACAGCAGCAGACACTTCAAGAAAGCTGCTGCTGCTGCTTAAAGGCTGAC 1620
Db 1719 CTCTCCACAGCAGCAGACACTTCAAGAAAGCTGCTGCTGCTTAAAGGCTGAC 1778
QY 1621 TGGATATCTTGGCCCTCAAGCTTCAACCTTTCCTGACAGCAGCTGTAATTTTGGC 1680
Db 1779 TGGATATCTTGGCCCTCAAGCTTCAACCTTTCCTGACAGCAGCTGTAATTTTGGC 1838
QY 1681 ACAGAGTCTCAGACTTCTACTGCTCAGGCACTTCAAGAGCAGTGCAGTTACTCT 1740

Db 1839 ACAGAGCCCTCACACTTCTACTCTGCTCAGCCATCTACGAGGCGATGCGCACTTCT 1898
Qy 1741 GCGTACAGCTGACGCCAGCTGCCCACTTGGCGAGACCAAGTCTATTCTGTGGCAGCGG 1800
Db 1899 GCGTACAGCTGACGCCAGCTGCCCACTTGGCGAGACCAAGTCTATTCTGTGGCAGCGG 1958
Qy 1801 CAGAGCGCAAGTACAGAGCTGACTCGCGGCGAGCTGGCATGAAGAGACCCCTTTGAA 1860
Db 1959 CAGAGCGCAAGTACAGAGCTGACTCGCGGCGAGCTGGCATGAAGAGACCCCTTTGAA 2018
Qy 1861 AACGAGTTTAAAGCAGAGAGCTGCCAATGGAATTTGGAGAGAGCATCATGTCAAGAAC 1920
Db 2019 AACGAGTTTAAAGCAGAGAGCTGCCAATGGAATTTGGAGAGAGCATCATGTCAAGAAC 2078
Qy 1921 AGGTACGCGGAGAGAGCTGGGGAAGTGGCGAGTCACTTCTTTCGGCGAGCATGAA 1980
Db 2079 AGGTACGCGGAGAGAGCTGGGGAAGTGGCGAGTCACTTCTTTCGGCGAGCATGAA 2138
Qy 1981 ATCATGAGGTCTCC 1995
Db 2139 ATCATGAGGTCTCC 2153

RESULT 15

AX713989

LOCUS AX713989 2102 bp DNA linear PART 15-APR-2003

DEFINITION Sequence 673 from Patent EPI293569.

ACCESSION AX713989

VERSION AX713989.1 GI:29888917

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 Isogai,T., Sugiyama,T., Otsuki,T., Makamatsu,A., Sato,H., Ishii,S.,

Yamanoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,

Tamehika,I., Seki,N., Yoshikawa,T., Otsuka,W., Nagaharti,K. and

Masuno,Y.

Full-length cDNA

Patent: EP 1293569-A 673 19-MAR-2003;

Helix Research Institute (JPI) ; Research Association for

Biotechnology (JPI)

Location/Qualifiers

1. 2102

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 99.7%; Score 1988.6; DB 6; Length 2102;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATTTGAACTCAATTTGTTACTGAGAGGTTGGTCTGCTGCGAA 60
Db 56 ATGGCCCATGAGATGATTTGAACTCAATTTGTTACTGAGAGGTTGGTCTGCTGCGAA 115
Qy 61 AGTGAAGCGAAAGAGTGTCTAATTTGATGAGCGGCGCATTTGGAATCAATACATCC 120
Db 116 AGTGAAGCGAAAGAGTGTCTAATTTGATGAGCGGCGCATTTGGAATCAATACATCC 175
Qy 121 CACATTTGGAGCCATTATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 180
Db 176 CACATTTGGAGCCATTATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 235
Qy 181 GACAAAGCTTTATATACAGAGCTATCCAGCATTTGAGCGAAACATAAGTTGACATTGAT 240
Db 236 GACAAAGCTTTATATACAGAGCTATCCAGCATTTGAGCGAAACATAAGTTGACATTGAT 295
Qy 241 TGCAGTCAGAAAGTTAGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 300

Db 296 TGCAGTCAGAAAGTTAGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 355
Qy 301 GACTGTTTCTCAGTGAATTTCTGGGTAACTGGAAGAGAGCTTCAACTTGTTCACCTG 360
Db 356 GACTGTTTCTCAGTGAATTTCTGGGTAACTGGAAGAGAGCTTCAACTTGTTCACCTG 415
Qy 361 CTTCGAGAGTGGGTTGCTGAGTCTCTCGTGTTCCTCGGAGCTCTGTAAGAAATCC 420
Db 416 CTTCGAGAGTGGGTTGCTGAGTCTCTCGTGTTCCTCGGAGCTCTGTAAGAAATCC 475
Qy 421 ACTCTAGTCCCTTACCTGATTTCTCAGCCCTGCTTACTGTTGCAACATTTGGGCCAAC 480
Db 476 ACTCTAGTCCCTTACCTGATTTCTCAGCCCTGCTTACTGTTGCAACATTTGGGCCAAC 535
Qy 481 CGAATCTTCCCAATCTTTATCTTGGCTGCGAGGAGATGTCTCAACAGAGCTGATA 540
Db 536 CTAAATCTTCCCAATCTTTATCTTGGCTGCGAGGAGATGTCTCAACAGAGCTGATA 595
Qy 541 CAGCAGAAATGGGATGGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCCTGA 600
Db 596 CAGCAGAAATGGGATGGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCCTGA 655
Qy 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGCTGATGAGAGCTTTGTGAGAAATTTTG 660
Db 656 ATCCCGAGTCTCATTTCTGCGTGTGCTGCTGATGAGAGCTTTGTGAGAAATTTTG 715
Qy 661 CCGTGGTGAACAATCAGTGAATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTGT 720
Db 716 CCGTGGTGGACAACAATCAGTGAATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTGT 775
Qy 721 CTAGTGCACGTGTTTACTGGGATCTCCCGCTCGGACCAATGCTTATGCTTATCATCAG 780
Db 776 CTAGTGCACGTGTTTACTGGGATCTCCCGCTCGGACCAATGCTTATGCTTATCATCAG 835
Qy 781 AAGAGAGTGAACATGCTTTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
Db 836 AAGAGAGTGAACATGCTTTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 895
Qy 841 ATATCTCCAAACTTCAATTTTCTGGGCGCACTCTGAGCTATGAGAAAGATTAAAGAAC 900
Db 896 ATATCTCCAAACTTCAATTTTCTGGGCGCACTCTGAGCTATGAGAAAGATTAAAGAAC 955
Qy 901 CAGACTGGAGCATCAGGGGCGAAAGAGCAAACTCAAGCTCTGACCTGGAGAAAGCCAAAT 960
Db 956 CAGACTGGAGCATCAGGGGCGAAAGAGCAAACTCAAGCTCTGACCTGGAGAAAGCCAAAT 1015
Qy 961 GAACCTGTCTCTGCTGTCTCAGAGGGTGGACAGAAAGCGAGAGCGCCCTCAGTCCACCC 1020
Db 1016 GAACCTGTCTCTGCTGTCTCAGAGGGTGGACAGAAAGCGAGAGCGCCCTCAGTCCACCC 1075
Qy 1021 TGTGCGACTGTGCTACTCAGAGGCGAGAGCAAAAGGCCGTGATCCCGCAGCGTG 1080
Db 1076 TGTGCGCGACTGTGCTACTCAGAGGCGAGAGCAAAAGGCCGTGATCCCGCAGCGTG 1135
Qy 1081 CCCAGGTCGCCAGCTGAGAGCGCTGTGTTAGAGGACAGAGCCCGTGTGTAAGAGGCGTC 1140
Db 1136 CCCAGGTCGCCAGCTGAGAGCGCTGTGTTAGAGGACAGAGCCCGTGTGTAAGAGGCGTC 1195
Qy 1141 AGTGGGCTGACCTGTGCGCGCAGAGAGGCTGGAAGACAGCAATTAAGCTCAAGCGTTC 1200
Db 1196 AGTGGGCTGACCTGTGCGCGCAGAGAGGCTGGAAGACAGCAATTAAGCTCAAGCGTTC 1255
Qy 1201 TCTTGGATATCAATCAATTTTATATTCAGCGAGATGGAGAGCTTCACTGAGCTTC 1260
Db 1256 TCTTGGATATCAATCAATTTTATATTCAGCGAGATGGAGAGCTTCACTGAGCTTC 1315
Qy 1261 TCTTATGAGAAAGTCTTGAATATCTCAAAAGCTTCACTGAGTGGAGAGCAAC 1320
Db 1316 TCTTATGAGAAAGTCTTGAATATCTCAAAAGCTTCACTGAGTGGAGAGCAAC 1375
Qy 1321 AAGCTATGCAAGTCTCCCTGTTCAGAACTATGAGAGAGAGCTCCGAAACAGTCTCT 1380
Db 1376 AAGCTATGCAAGTCTCCCTGTTCAGAACTATGAGAGAGAGCTCCGAAACAGTCTCT 1435

```
QY 1381 GATAGAGAGAGAGAGATCCCAAGAGCTGAGACCGCCAGGCGCTTCAGACAGCCAG 1440
Db 1436 GATAGAGAGAGAGAGAGATCCCAAGAGCTGAGACCGCCAGGCGCTTCAGACAGCCAG 1495
QY 1441 AGCAAGCGATTGCAATTGGGTCAAGAACAGAGAGTGGCAACCGCCAGAGGTCCTTTTA 1500
Db 1496 AGCAAGCGATTGCAATTGGGTCAAGAACAGAGAGTGGCAACCGCCAGAGGTCCTTTTA 1555
QY 1501 TCTCACTGATCGAAGTGGAGAGCGTGGAGCAATTACACACAGCTTCCCTTTGGGC 1560
Db 1556 TCTCACTGATCGAAGTGGAGAGCGTGGAGCAATTACACACAGCTTCCCTTTGGGC 1615
QY 1561 CTTTCCACAGCCAGAGCACTTCACGAGTCTGCTGGCTGGGCTTAAGGGCTGGCAC 1620
Db 1616 CTTTCCACAGCCAGAGCACTTCACGAGTCTGCTGGCTGGGCTTAAGGGCTGGCAC 1675
QY 1621 TCGGATATCTTGGGCCCCCGAGACCTCTACCCCTTCCCTGACACAGAGCTGTATTTTGGC 1680
Db 1676 TCGGATATCTTGGGCCCCCGAGACCTCTACCCCTTCCCTGACACAGAGCTGTATTTTGGC 1735
QY 1681 ACAGAGTCTTCACACTTCTACTCTGCTGAGCCATCTACGAGAGGAGTGCAGTTACTCT 1740
Db 1736 ACAGAGTCTTCACACTTCTACTCTGCTGAGCCATCTACGAGAGGAGTGCAGTTACTCT 1795
QY 1741 GCCTACAGCTGCAAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGCGCAGGCGG 1800
Db 1796 GCCTACAGCTGCAAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGCGCAGGCGG 1855
QY 1801 CAGAGCCAAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATGAGAGAGCCCTTTGAA 1860
Db 1856 CAGAGCCAAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATGAGAGAGCCCTTTGAA 1915
QY 1861 AAGCACTTTAAAGCGCAAGCTGCCAAATGAAATTTGAGAGAGCATCATGTCAAGAAAC 1920
Db 1916 AAGCACTTTAAAGCGCAAGCTGCCAAATGAAATTTGAGAGAGCATCATGTCAAGAAAC 1975
QY 1921 AGGTACGGGAAAGAGCTGGGAAAGTGGGCACTCACTTCTAGTTTTCGGGCAAGTGGAA 1980
Db 1976 AGGTACGGGAAAGAGCTGGGAAAGTGGGCACTCACTTCTAGTTTTCGGGCAAGTGGAA 2035
QY 1981 ATCATTTGAGGTCCTCC 1995
Db 2036 ATCATTTGAGGTCCTCC 2050
```

Search completed: June 21, 2004, 23:40:19
Job time : 12565.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 17:00:31 ; Search time 7898.16 Seconds

(Without alignments)

7542.898 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995

Sequence: 1 atggcccatgagatgatg.....tggaatcattgaggtctcc 1995

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estbm:*
3: em_estbn:*
4: em_estbu:*
5: em_estbv:*
6: em_estbl:*
7: em_estro:*
8: em_estc:*
9: gb_estc1:*
10: gb_estc2:*
11: gb_hnc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estbm:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_juv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rnd:*
26: em_gss_png:*
27: em_gss_vrl:*
28: gb_gss2:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	896.8	45.0	1060	12	BM546940 AGENCOURT
2	834.6	41.8	883	13	BQ945892 AGENCOURT
3	827.8	41.5	881	13	BQ930140 AGENCOURT
4	809.6	40.6	859	13	BQ933499 AGENCOURT

Result No.	Score	Query Match	Length	ID	Description
5	799.8	40.1	898	13	BQ721265 AGENCOURT
6	786.2	39.4	974	13	BQ951695 AGENCOURT
7	780.2	39.1	903	12	BG482429 AGENCOURT
8	695.6	34.9	836	12	B1821804 AGENCOURT
9	660.6	33.1	862	13	BQ770036 AGENCOURT
10	641.2	32.1	920	10	BE8897795 AGENCOURT
11	633	31.7	682	12	B1816954 AGENCOURT
12	627	31.4	663	13	BK479029 AGENCOURT
13	623.2	31.2	891	13	BQ670889 AGENCOURT
14	615.4	30.8	769	13	BU704078 AGENCOURT
15	606.6	30.4	983	12	B165188 AGENCOURT
16	603.2	30.2	752	13	BQ432082 AGENCOURT
17	592.6	29.7	715	14	CF742387 AGENCOURT
18	559.6	28.1	639	14	CF742387 AGENCOURT
19	558.2	28.0	656	14	CF532917 AGENCOURT
20	544.8	27.3	650	10	BF472046 AGENCOURT
21	541.4	27.1	602	10	AM847425 AGENCOURT
22	541.2	27.1	621	9	AV703072 AGENCOURT
23	541.2	26.9	693	9	AV701628 AGENCOURT
24	536.8	26.8	595	10	AM847426 AGENCOURT
25	527.2	26.4	790	12	B1661614 AGENCOURT
26	521	26.1	946	12	B1657528 AGENCOURT
27	490.8	24.6	743	14	CF727802 AGENCOURT
28	482.8	24.2	3325	11	AK035652 AGENCOURT
29	479.4	24.0	739	28	AZ850283 AGENCOURT
30	462.4	23.8	631	28	BH039241 AGENCOURT
31	455.4	22.8	1067	10	BF135687 AGENCOURT
32	455.4	22.8	792	13	BU946569 AGENCOURT
33	428.8	21.5	512	10	BF815601 AGENCOURT
34	425.8	21.3	586	14	CF104291 AGENCOURT
35	422.8	21.2	516	29	AY413655 AGENCOURT
36	414	20.8	723	13	BU708175 AGENCOURT
37	404.6	20.3	619	29	CG583869 AGENCOURT
38	399.6	20.0	463	10	BG058779 AGENCOURT
39	389.4	19.5	391	13	BK479224 AGENCOURT
40	386.4	19.4	463	10	BF809325 AGENCOURT
41	386.2	19.4	2207	11	BC038231 AGENCOURT
42	379	19.0	444	12	BM987142 AGENCOURT
43	373.2	18.7	516	29	AY413657 AGENCOURT
44	370.6	18.6	594	14	CF727694 AGENCOURT
45					

ALIGNMENTS

RESULT 1
BM546940
LOCUS
DEFINITION
AGENCOURT_6491295 NIH_MGC_125 Homo sapiens cdna clone IMAGE:5723711
5', mRNA sequence.
BM546940
AGENCOURT_6491295 NIH_MGC_125 Homo sapiens cdna clone IMAGE:5723711
BM546940.1 GI:18780310
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
TITLER
JOURNAL
COMMENT
NIH-MGC http://img.ncbi.nlm.nih.gov/
1 (bases 1 to 1060)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LLM12711 row: m column: 24
High quality sequence spot: 737.

FEATURES

source

Location/Qualifiers

```

1..1060
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5723711"
/lab_host="DH10B"
/clone.lib="NIH_MGC_125"
/Note="Organ: ovary (pool of 3); Vector: pCMV-Sport6;
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

```

ORIGIN

Query Match 45.0%; Score 896.8; DB 12; Length 1060;
 Best Local Similarity 94.7%; Pred. No. 2,9e-241;
 Matches 973; Conservative 0; Mismatches 37; Indels 17; Gaps 4;

```

QY 488 TTCCCAATCTTATCTGCTGCGACGAGATGCTCTCAACAGAGCTGATACAGAGA 547
Db 7 TTCCCAATCTTATCTGCTGCGACGAGATGCTCTCAACAGAGCTGATACAGAGA 66
QY 548 ATGGAGATGGTATGCTTAAATGCGAGTATACCTGTCAAAGCTGACCTTATCCCG 607
Db 67 ATGGAGATGGTATGCTTAAATGCGAGTATACCTGTCAAAGCTGACCTTATCCCG 126
QY 608 AGTCTCATTTCTGCGTGTGCTGCTGAGTACAGCTTTGTGAGAAATTTTCCGTGT 667
Db 127 AGTCTCATTTCTGCGTGTGCTGCTGAGTACAGCTTTGTGAGAAATTTTCCGTGT 186
QY 668 TGGACAAATCAGTATATTTCAATGAGAAAGCAAAAGCCCTCAATGATGTTTAAAGC 727
Db 187 TGGACAAATCAGTATATTTCAATGAGAAAGCAAAAGCCCTCAATGATGTTTAAAGC 246
QY 728 ACTGTTTATGCTGGATCTCCGCTCCGCAACATCCGCTATCCGCTATCATATGAAAGA 787
Db 247 ACTGTTTATGCTGGATCTCCGCTCCGCAACATCCGCTATCCGCTATCATATGAAAGA 306
QY 788 TGGACATGCTTTTATGAGAGCTTACATGATTGTGAAAGAAAAAGACCTATATATCTC 847
Db 307 TGGACATGCTTTTATGAGAGCTTACATGATTGTGAAAGAAAAAGACCTATATATCTC 366
QY 848 CAAACTTCAATTTTCTGGGCAACTCTGAGATGAGAGAAAGTTTAAAGCAAGCATG 907
Db 367 CAAACTTCAATTTTCTGGGCAACTCTGAGATGAGAGAAAGTTTAAAGCAAGCATG 426
QY 908 GAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTTGAGAGAGCCAAATGAACTG 967
Db 427 GAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTTGAGAGAGCCAAATGAACTG 486
QY 968 TCCCTGCTGTCTGAGAGGCTGAGAGAAAGAGCCGCTCAGTCCACCTGTGCGG 1027
Db 487 TCCCTGCTGTCTGAGAGGCTGAGAGAAAGAGCCGCTCAGTCCACCTGTGCGG 546
QY 1028 ACTGCTACTACCTCAGAGGAGAGAGAAAGCCCGTCAATCCCGCAGCTGCGCAGG 1087
Db 547 ACTGCTACTACCTCAGAGGAGAGAGAAAGCCCGTCAATCCCGCAGCTGCGCAGG 1087
QY 1088 TGGCCAGGCTGAGAGCCGCTGCTGTTAGAGAGAGCCGCTGAGTCAAGGGCTCACTGAGG 1147
Db 598 TGGCCAGGCTGAGAGCCGCTGCTGTTAGAGAGAGCCGCTGAGTCAAGGGCTCACTGAGG 657
QY 1148 TGCACCTGTCCGAGAGGCTGAGAGAGAGCAAGCAATTAAGCTCAAGGCTTCTCTCTG 1207
Db 658 TGCACCTGTCCGAGAGGCTGAGAGAGAGCAAGCAATTAAGCTCAAGGCTTCTCTCTG 717
QY 1208 ATATCAATATAGTTTCAATATTAAGCAGATGAGAGCATCTCTCAAGGCTTCTCTCAT 1267

```

```

Db 718 ATATCAATATAGTTTCAATATTAAGCAGATGAGAGCATCTCTCAAGGCTTCTCTCAT 777
QY 1268 CAGAAAGTGGTGTGAATATCTAACAAACCTTCCACTACTGTGAGAGGAGCAACAGCTTAT 1327
Db 778 CAGAAAGTGGTGTGAATATCTAACAAACCTTCCACTACTGTGAGAGGAGCAACAGCTTAT 837
QY 1328 GCCAGTTCTCCCTGTGAGAGAGCAATTAAGGAGCAAGTCCCGAAACCAAGTCTGATTAAG 1387
Db 838 GCCAGTTCTCCCTGTGAGAGAGCAATTAAGGAGCAAGTCCCGAAACCAAGTCTGATTAAG 896
QY 1388 AGGAAGCCAGCATCTCCCAAGAGCTGAGAGCCGCA-GGCTTCAAGAGCCAGAGCAAG 1446
Db 897 AGGAAGCCAGCATCTCCCAAGAGCTGAGAGCCGCAAGGCTTCAAGAGCCAGAGCAAG 956
QY 1447 CG-----ATTGCAATTCGCTGAGAGAGCAAGAGTGGCAACCGCCAGAGGTCCTTTTA 1500
Db 957 CGAATTCATTCCTGAGAGAGCAAGAGCAAGAGTGGCAACCGCCAGAGGTCCTTTTA 1501
QY 1501 TCTCCAC 1507
Db 1017 TATCCTC 1023

```

RESULT 2
 B0945892
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 883)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNI)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNI at:
 http://image.llnl.gov
 Plate: LHC2653 row: o column: 14
 High quality sequence stop: 672.

FEATURES

source

Location/Qualifiers

```

1..883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6463045"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone.lib="NIH_MGC_101"
/Note="Organ: lung; Vector: pOT87; Site 1: EcoRI; Site 2:
XhoI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

ORIGIN

Query Match 41.8%; Score 834.6; DB 13; Length 883;
 Best Local Similarity 98.5%; Pred. No. 9,4e-224;
 Matches 874; Conservative 0; Mismatches 9; Indels 4; Gaps 3;


```

QY 794 TGTCTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGACCTACTATATCTCCAACT 853
DB 1 TGTCTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGACCTACTATATCTCCAA--C 57
QY 854 TCAATTTTCTGGGCAAACTCTGGACTATGAGAAGATTAAAGAACCAAGCTGGAGCAT 913
DB 58 TTGATTTTCTGGGCAAACTCTGGACTATGAGAAGATTAAAGAACCAAGCTGGAGCAT 117
QY 914 CAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAACCAATGAATCTGTCCCTG 973
DB 118 CAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAACCAATGAATGAATCTGTCCCTG 177
QY 974 CTGTCTCAGAGGGGTGACAGAAAAGGAGAGGCCCTCAGTCCACCTGTGCGGACTGTG 1033
DB 178 CTGTCTCAGAGGGGTGACAGAAAAGGAGAGGCCCTCAGTCCACCTGTGCGGACTGTG 237
QY 1034 CTACCTCAGAGGAGAGAGCAAAAGGCCCTGTCATCCCGCAGAGGTGCCAGGCTGCCA 1093
DB 238 CTACCTCAGAGGAGAGAGCAAAAGGCCCTGTCATCCCGCAGAGGTGCCAGGCTGCCA 297
QY 1094 GCGTGACCGCTGCTGTGTAAGAGACAGCCCGCTGTGTAAGAGGCTCAGTGGGCTGCACC 1153
DB 298 GCGTGACCGCTGCTGTGTAAGAGACAGCCCGCTGTGTAAGAGGCTCAGTGGGCTGCACC 357
QY 1154 TGTCCGACAGAGGCTGGAGAGACAGCAATAAGCTCAAGCGTTCTTCTCTGTGATATCA 1213
DB 358 TGTCCGACAGAGGCTGGAGAGACAGCAATAAGCTCAAGCGTTCTTCTCTGTGATATCA 417
QY 1214 AATCAGTTTCAATTTAGCAGAGATGAGCAACATCTTAACATGAGTTCTCTCATCAGAA 1273
DB 418 AATCAGTTTCAATTTAGCAGAGATGAGCAACATCTTAACATGAGTTCTCTCATCAGAA 477
QY 1274 ATGCTTTGAATATCTACAACTTTCACATCTCTGATGAGGAGCAACAGCTATGCGAGT 1333
DB 478 ATGCTTTGAATATCTACAACTTTCACATCTCTGATGAGGAGCAACAGCTATGCGAGT 537
QY 1334 TTTCCCTGTTTCAAGAACTATCGAGAGCACTCCCGAAACAGTCCGTGATAGAGAGAG 1393
DB 538 TTTCCCTGTTTCAAGAACTATCGAGAGCACTCCCGAAACAGTCCGTGATAGAGAGAG 597
QY 1394 CAGAGATCCCAAGAAAGCTGACAGACCGCCAGGCTTTCAGACAGCAGAGCAAGGATTCG 1453
DB 598 CAGAGATCCCAAGAAAGCTGACAGACCGCCAGGCTTTCAGACAGCAGAGCAAGGATTCG 657
QY 1454 ATTGCTGACAGAACCAAGAGAGAGTGGACCGCCAGAGGCTCTTTATCTCTCACTGATC 1513
DB 658 ATTGCTGACAGAACCAAGAGAGAGTGGACCGCCAGAGGCTCTTTATCTCTCACTGATC 717
QY 1514 GAAGTGGAGAGCTGGAGAGCAATTAACAACACAGCTTCTTTTGGGCTTTCCACCGGCC 1573
DB 718 GAAGTGGAGAGCTGGAGAGCAATTAACAACACAGCTTCTTTTGGGCTTTCCACCGGCC 777
QY 1574 AGCAGCACTCAAGAGTCTGCTGAGCTGGAGCTTAAAGGCTGGCACTCGGATATCTTGG 1633
DB 778 AGCAGCACTCAAGAGTCTGCTGAGCTGGAGCTTAAAGGCTGGCACTCGGATATCTTGG 836
QY 1634 CCCCCCAGAGCTTACCTCTTCCCTTCAACAGCAGCTGTATTTTGGC 1680
DB 837 CCCCCCAGAGCTTACCTCTTCCCTTCAACAGCAGCTGTATTTTGGC 883

```

```

RESULT 3
BQ930140 881 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT 8923732 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6460964
DEFINITION 5', mRNA sequence.
ACCESSION BQ930140
VERSION BQ930140.1 GI:22345171
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC http://imgc.ncl.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LCM2648 row: h column: 21
High quality sequence step: 676.
Location/Qualifiers
1. 881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6460964"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/note="Organ: Lung; Vector: POTB1; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming, directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

ORIGIN

```

Query Match 41.5%; Score 827.8; DB 13; Length 881;
Best Local Similarity 99.4%; Pred. No. 7.8e-222;
Matches 862; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 794 TGTCTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGACCTACTATATCTCCAACT 853
DB 1 TGTCTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGACCTACTATATCTCCAACT 60
QY 854 TCAATTTTCTGGGCAAACTCTGGACTATGAGAAGATTAAAGAACCAAGCTGGAGCAT 913
DB 61 TCAATTTTCTGGGCAAACTCTGGACTATGAGAAGATTAAAGAACCAAGCTGGAGCAT 120
QY 914 CAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAACCAATGAATGAATCTGTCCCTG 973
DB 121 CAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAACCAATGAATGAATCTGTCCCTG 180
QY 974 CTGTCTCAGAGGGGTGACAGAAAAGGAGAGCGCCCTCAGTCCACCTGTGCGGACTGTG 1033
DB 181 CTGTCTCAGAGGGGTGACAGAAAAGGAGAGCGCCCTCAGTCCACCTGTGCGGACTGTG 240
QY 1034 CTACCTCAGAGGAGAGAGCAAAAGGCCGTGATCCCGCAGAGGTGCCAGGCTGCCA 1093
DB 241 CTACCTCAGAGGAGAGAGCAAAAGGCCGTGATCCCGCAGAGGTGCCAGGCTGCCA 300
QY 1094 GCGTGACCGCTGCTGTGTAAGAGACAGCCCGCTGTGTAAGAGGCTCAGTGGGCTGCACC 1153
DB 301 GCGTGACCGCTGCTGTGTAAGAGACAGCCCGCTGTGTAAGAGGCTCAGTGGGCTGCACC 360
QY 1154 TGTCCGACAGAGGCTGGAGAGACAGCAATAAGCTCAAGCGTTCTTCTCTGTGATATCA 1213
DB 361 TGTCCGACAGAGGCTGGAGAGACAGCAATAAGCTCAAGCGTTCTTCTCTGTGATATCA 420
QY 1214 AATCAGTTTCAATTTAGCAGAGATGAGCAACATCTTAACATGAGTTCTCTCATCAGAA 1273
DB 421 AATCAGTTTCAATTTAGCAGAGATGAGCAACATCTTAACATGAGTTCTCTCATCAGAA 480
QY 1274 ATGCTTTGAATATCTACAACTTTCACATCTCTGATGAGGAGCAACAGCTATGCGAGT 1333
DB 481 ATGCTTTGAATATCTACAACTTTCACATCTCTGATGAGGAGCAACAGCTATGCGAGT 540

```

QY 1334 TCTCCCTGTGAGAGACTATGAGAGAGACTCCGAAACAGTCTGATAGAGAGAG 1393
Db 541 TCTCCCTGTGAGAGACTATGAGAGAGACTCCGAAACAGTCTGATAGAGAGAG 600
QY 1394 CAGAGATCCCAAGAGAGTGCAGACCGCCAGGCTTTCAGACAGCAGAGAGAGATTGC 1453
Db 601 CCAGATCCCAAGAGAGTGCAGACCGCCAGGCTTTCAGACAGCAGAGAGAGATTGC 660
QY 1454 ATTGCGTGAAGACAGAGAGTGCAGACCGCCAGGCTTTCAGACAGCAGAGAGATTGC 1513
Db 661 ATTGCGTGAAGACAGAGAGTGCAGACCGCCAGGCTTTCAGACAGCAGAGAGATTGC 720
QY 1514 GAAGTGGAGGCTGAGAGAGATTAACAGACAGCTTCC-TTTTGGGCTTTTCACAGC 1572
Db 721 GAAGTGGAGGCTGAGAGAGATTAACAGACAGCTTCC-TTTTGGGCTTTTCACAGC 780
QY 1573 CAGAGAGACCTCAAGAGTCTGAGGCTT-GGGCTTTAAGGAGTGCAGTGCAGATATCTT 1631
Db 781 CAGAGAGACCTCAAGAGTCTGAGGCTTTCAGGAGCTTAAAGGCTGAGCTGCAGATATCTT 840
QY 1632 GG-CCCCCAGACCTTACCCCTTCC 1657
Db 841 GGCCCCCAGACCTTACCCCTTCC 867

RESULT 4
B0933499 859 bp mRNA linear EST 21-ANG-2002
LOCUS AGENCOURT 8732489 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6455595
DEFINITION 5', mRNA sequence.
ACCESSION B0933499
VERSION B0933499.1 GI:22348882
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 859)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM2634 row: 1 column: 04
High quality sequence stop: 727.
Location/Qualifiers
1. 859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="IMAGE:6455595"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_101"
/note="Organ: lung; Vector: POTB1; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN
Query Match 40.6%; Score 809.6; DB 13; Length 859;

Best Local Similarity 98.5%; Pred. No. 1,1e-216;
Matches 827; Conservative 0; Mismatches 11; Indels 2; Gaps 1;
QY 794 TGTCTTTAATGATGAGCTTAAGATTGTGAAAGAAAGAAAGCACTATATCTTCAACT 853
Db 1 TGTCTTTAATGATGAGCTTAAGATTGTGAAAGAAAGAAAGCACTATATCTTCAACT 60
QY 854 TCAATTTTGTGGGCCAATCTCCCTGATCTTGAAGAAATTAAGAACAGACTGAGCAT 913
Db 61 TCAATTTTGTGGGCCAATCTCCCTGATCTTGAAGAAATTAAGAACAGACTGAGCAT 120
QY 914 CAGGGCCAAAGAGCAAACTCAAGCTGTGACCTGTGAGAAAGCAAAATGAACCTGTCTCTG 973
Db 121 CAGGGCCAAAGAGCAAACTCAAGCTGTGACCTGTGAGAAAGCAAAATGAACCTGTCTCTG 180
QY 974 CTGTCTCAAGAGGAGTGAAGAGAAAGAGAGAGCCCTTCACTCACTGCTGTGCTGTG 1033
Db 181 CTGTCTCAAGAGGAGTGAAGAGAAAGAGAGAGCCCTTCACTCACTGCTGTGCTGTG 240
QY 1034 CTACTCTCAAGAGGAGAGAGCAAAAGCCGTGATCCGCGAGGTCAGAGCTGAGCCCA 1093
Db 241 CTACTCTCAAGAGGAGAGAGCAAAAGCCGTGATCCGCGAGGTCAGAGCTGAGCCCA 300
QY 1094 GCGTGCAGCCGTGCTGTTAAGAGAGAGCCGCTGTACAGGCTCAAGTGCAGTGCACC 1153
Db 301 GCGTGCAGCCGTGCTGTTAAGAGAGAGCCGCTGTACAGGCTCAAGTGCAGTGCACC 360
QY 1154 TGTGCCAGACAGGCTGTGAAGACAGCAATTAAGCTCAAGCTTCTCTGTGATATCA 1213
Db 361 TGTGCCAGACAGGCTGTGAAGACAGCAATTAAGCTCAAGCTTCTCTGTGATATCA 420
QY 1214 AATCAGTTTCATATTCAGCAGCAATGAGAGATCTTACATGCTTCTCTCATAGAG 1273
Db 421 AATCAGTTTCATATTCAGCAGCAATGAGAGATCTTACATGCTTCTCTCATAGAG 480
QY 1274 ATGCTTTGGAATACATCAAAACCTTCACTCACTGTGAGAGCAAAAGCTATGCACT 1333
Db 481 ATGCTTTGGAATACATCAAAACCTTCACTCACTGTGAGAGCAAAAGCTATGCACT 540
QY 1334 TCTCCCTGTTCAGGAACTATGAGAGAGAGCTCCGAAACAGTCTGATAGAGAGAG 1393
Db 541 TCTCCCTGTTCAGGAACTATGAGAGAGAGCTCCGAAACAGTCTGATAGAGAGAG 600
QY 1394 CAGAGATCCCAAGAGAGTGCAGACCGCCAGGCTTTCAGACAGCAGAGAGAGATTGC 1453
Db 601 CAGAGATCCCAAGAGAGTGCAGACCGCCAGGCTTTCAGACAGCAGAGAGAGATTGC 660
QY 1454 ATTGCGTGAAGACAGAGAGTGCAGACCGCCAGGCTTTCAGACAGCAGAGAGATTGC 1513
Db 661 ATTGCGTGAAGACAGAGAGTGCAGACCGCCAGGCTTTCAGACAGCAGAGAGATTGC 720
QY 1514 GAAGTGGAGGCTGAGAGAGATTAACAGACAGCTTCC-TTTTGGGCTTTTCACAGC 1572
Db 721 GAAGTGGAGGCTGAGAGAGATTAACAGACAGCTTCC-TTTTGGGCTTTTCACAGC 780
QY 1574 AGCAGACCTCAAGAGTCTGAGGCTT-TGGGCTTTAAGGCTGAGCTGCAGATATCTT 1631
Db 781 AGCAGACCTCAAGAGTCTGAGGCTT-TGGGCTTTAAGGCTGAGCTGCAGATATCTT 840

RESULT 5
B0721265 898 bp mRNA linear EST 16-JUL-2002
LOCUS AGENCOURT 8294508 Lupski symmetric_trunk Homo sapiens cDNA clone
DEFINITION IMAGE:6194455 5', mRNA sequence.
ACCESSION B0721265
VERSION B0721265.1 GI:21860162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 898)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1AM13599 row: 1 column: 08
 High quality sequence stop: 669.
 Location/Qualifiers

FEATURES

source

1. 898
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6194455"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /clone_1lb="Lupski, sympathetic trunk"
 /note="Vector: pCMV-Sport6 (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCACCGCTCCG-3' and
 5'-GACTAGTTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected >
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine); available through Life
 Technologies."

ORIGIN

Query Match 40.1%; Score 799.8; DB 13; Length 898;
 Best Local Similarity 96.8%; Pred. No. 6,4e-214;
 Matches 859; Conservative 0; Mismatches 22; Indels 6; Gaps 4;

QY 8 ATGAGATGATGGAACTCAATTGTACTGAGAGGTGTGGCTGTGCTGGAAGTGAA 67
 DB 12 ATGAGATGATGGAACTCAATTGTACTGAGAGGTGTGGCTGTGCTGGAAGTGAA 71
 QY 68 CGAAGAAAGTGTCTAATTGATAGCGGCGCATTTGTGAATACAAATACATCCCATTT 127
 DB 72 CGAAGAAAGTGTCTAATTGATAGCGGCGCATTTGTGAATACAAATACATCCCATTT 131
 QY 128 TGAAGCCATTAATATGACTGCTCCAGGTTATGAAGCGAAGTTGGACAGACAAAG 187
 DB 132 TGAAGCCATTAATATGACTGCTCCAGGTTATGAAGCGAAGTTGGACAGACAAAG 191
 QY 188 TGTAAATACAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGATGGAGTC 247
 DB 192 TGTAAATACAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGATGGAGTC 251
 QY 248 AGAAGTTGATTTAGATCAAGCTCCCAAGATGTGCTCTCTCTTCAGACTGTT 307
 DB 252 AGAAGTTGATTTAGATCAAGCTCCCAAGATGTGCTCTCTCTTCAGACTGTT 311
 QY 308 TTCTCATCTGATCTTGGGTTAACTGAGAGAGCTTCACTCTGTTCACTGCTGGAG 367
 DB 312 TTCTCATCTGATCTTGGGTTAACTGAGAGAGCTTCACTCTGTTCACTGCTGGAG 371
 QY 368 GTGGGTTTGTGAGTTCTCTGTTGTTCCCTGGCCTCTGTGAAGAGAAATCACTCTAG 427
 DB 372 GTGGGTTTGTGAGTTCTCTGTTGTTCCCTGGCCTCTGTGAAGAGAAATCACTCTAG 431
 QY 428 TCCCTACCTGCAATTTCTAGCCTTGTCTTACCTGTTGCAACATTTGGGCAACCCGAATTC 487
 DB 432 TCCCTACCTGCAATTTCTAGCCTTGTCTTACCTGTTGCAACATTTGGGCAACCCGAATTC 491

QY 488 TTCCCAATCTTTATCTTGCTGCGACGAGATGTCCTCAACAGAGCTGATACAGAGA 547
 DB 492 TTCCCAATCTTTATCTTGCTGCGACGAGATGTCCTCAACAGAGCTGATACAGAGA 551
 QY 548 ATGGAGTTGGTTATGTTGTTAAATGCGAGCTATACCTGTCCAAAGCTGACTTATCCCG 607
 DB 552 ATGGAGTTGGTTATGTTGTTAAATGCGAGCTATACCTGTCCAAAGCTGACTTATCCCG 611
 QY 608 AGCTCATTTTCCCTGCTGCTGCTGTGTAAGACAGCTTTTGTGAGAAATTTCCCGGT 667
 DB 612 AGCTCATTTTCCCTGCTGCTGCTGTGTAAGACAGCTTTTGTGAGAAATTTCCCGGT 670
 QY 668 TGGACAATCAGATGATTCATTTGATGAGAAAGCAAGCTTCATGATGATGTTCTAGTC 727
 DB 671 TGGACAATCAGATGATTCATTTGATGAGAAAGCAAGCTTCATGATGATGTTCTAGTC 730
 QY 728 ACTGTTTAACTGGATCTCCGCTCCGCAACATGCTATGCTTACATCATGTAAGAGA 787
 DB 731 ACTGTTTAACTGGATCTCCGCTCCGCAACATGCTATGCTTACATCATGTAAGAGA 789
 QY 788 TGGACATGTC-TTTAATGAGAGCTTACAGATTTGTGAAAGAAAAAGACTTATATCT 846
 DB 790 TGGACATGTC-TTTAATGAGAGCTTACAGATTTGTGAAAGAAAAAGACTTATATCT 849
 QY 847 CCAACTT---CAATTTCTGGGCAACTCTGGAATGAGAAAGA 890
 DB 850 CCAAACTTCAATTTCTTGGGCAACTCTGGAATGAGAAAGA 896

RESULT 6

B0951695

LOCUS

DEFINITION

AGENCOURT_9947130 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6461668

5', mRNA sequence.

B0951695

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 974)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: L1CM2650 row: f column: 05

High quality sequence stop: 617.

Location/Qualifiers

1. 974

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6461668"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_1lb="NIH_MGC_101"
 /note="Organ: lung; Vector: pOT87; Site 1: EcoRI; site 2:
 XhoI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Query Match 39.4%; Score 786.2; DB 13; Length 974;
 Best Local Similarity 94.5%; Pred. No. 4.7e-210;
 Matches 882; Conservative 0; Mismatches 38; Indels 13; Gaps 6;

QY 794 TGTCTTTAGATGAACTTACAGATTGTGAAAGAAAAAAGCTTACTATATCTCCAACT 853
 DB 1 TGTCTTTAGATGAACTTACAGATTGTGAAAGAAAAAAGCTTACTATATCTCCAACT 60
 QY 854 TCAATTTCTGGGGCCAACTCTGACATATGAGAAAGATTAAAGAACAGACTGAGCAT 913
 DB 61 TCAATTTCTGGGGCCAACTCTGACATATGAGAAAGATTAAAGAACAGACTGAGCAT 120
 QY 914 CAGGGCCAAAGACAACTCAAGCTGTGACCTGAGAGAGCCAAATGAACTGTCTCTG 973
 DB 121 CAGGGCCAAAGACAACTCAAGCTGTGACCTGAGAGAGCCAAATGAACTGTCTCTG 180
 QY 974 CTGTCTGAGAGGGTGGACAGAAAAAGGAGACCGCTCACTGACCTGTGCGCACTCTG 1033
 DB 181 CTGTCTGAGAGGGTGGACAGAAAAAGGAGACCGCTCACTGACCTGTGCGCACTCTG 240
 QY 1034 CTACCTTACAGGAGAGAGACAAAGGCGGTGATCCGCGCAGGCTGCGGCGGCTGCGCA 1093
 DB 241 CTACCTTACAGGAGAGAGACAAAGGCGGTGATCCGCGCAGGCTGCGGCGGCTGCGCA 300
 QY 1094 GCGTGGAGCGGTGCTGTTAGAGAGACCGCTGTGTAAGGCGCTCACTGAGGCTGCGCA 1153
 DB 301 GCGTGGAGCGGTGCTGTTAGAGAGACCGCTGTGTAAGGCGCTCACTGAGGCTGCGCA 360
 QY 1154 TGTCCGACACAGCGCTGAGAGACAGCAATTAAGCTCAAGCGTCTCTCTCTGAGATATCA 1213
 DB 361 TGTCCGACACAGCGCTGAGAGACAGCAATTAAGCTCAAGCGTCTCTCTCTGAGATATCA 420
 QY 1214 AATCAGTTTCATATTCAGCCAGCATGGAGATCTTAACTGTGCTTCTCTCTATCAGAA 1273
 DB 421 AATCAGTTTCATATTCAGCCAGCATGGAGATCTTAACTGTGCTTCTCTCTATCAGAA 480
 QY 1274 ATGCTTGGAAATCTACAACTTTCATCTCTGATGGAGACCAACAGCTATGCTCAGT 1333
 DB 481 ATGCTTGGAAATCTACAACTTTCATCTCTGATGGAGACCAACAGCTATGCTCAGT 540
 QY 1334 TCTCCCTGTTTCAAGAACTATCGAGAGCATCTCCGAAACCAAGTCTGTATTAAGAGAA 1393
 DB 541 TCTCCCTGTTTCAAGAACTATCGAGAGCATCTCCGAAACCAAGTCTGTATTAAGAGAA 600
 QY 1394 CCAGCATCCCAAGAAAGCTGACAGCCGCAAGGCTTCAAGACGACAGCAAGGAGATTGC 1453
 DB 601 CCAGCATCCCAAGAAAGCTGACAGCCGCAAGGCTTCAAGACGACAGCAAGGAGATTGC 660
 QY 1454 ATTGGTCAAGAACACAGACAGATGGACCGCCCAAGGTCCTTTATCTTCACTGCATC 1513
 DB 661 ATTGGTCAAGAACACAGACAGATGGACCGCCCAAGGTCCTTTATCTTCACTGCATC 720
 QY 1514 GAAGTGGAGCGGTGAGAGCAATTAACAACACAGCTT---CCTTTTGGGCTTTTCCACA 1570
 DB 721 GAAGTGGAGCGGTGAGAGCAATTAACAACACAGCTTCCCTTTTGGGCTTTTCCACA 780
 QY 1571 GCCAGACAGA-CCTTACGAAGTCTGCTGGCTGGG--CCTTAAAGGCTGACATCCGA-- 1625
 DB 781 GCCAGACAGACCTTACGAATCTGCTGGCTGGGCTTAAAGGCTGACATCCGAAT 840
 QY 1626 TATCTTGGCCCCCAGACCTTACCCCTT---CCCTGACAGACAGCTGTATTTTGCACA 1683
 DB 841 ATCTTTGGGGCCCCAGACCTTACCCCTTCCCTGAAACAAAGAGCTGTATTTTGCAC 900
 QY 1684 GAGTCTTACA---CTTCTACTCTGCTGAGCC 1713
 DB 901 GAAAGCCCTCACACCTTCTACTCTGCTGAGCC 933

RESULT 7
 BG482429

LOCUS BG482429 903 bp mRNA linear EST 21-MAR-2001
 DEFINITION 602526595F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4650430 5',
 mRNA sequence.
 ACCESSION BG482429
 VERSION BG482429.1 GI:13414708
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 903)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rewmail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.jhmi.gov
 Plate: L10M1431 row: m column: 23
 High quality sequence stop: 806.
 Location/Qualifiers
 1. 903
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4650430"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: placenta; Vector: pOT7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACAG(G). Size-selected
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

FEATURES

source
 1. 903
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4650430"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: placenta; Vector: pOT7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACAG(G). Size-selected
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 39.1%; Score 780.2; DB 12; Length 903;
 Best Local Similarity 94.7%; Pred. No. 2.2e-208;
 Matches 853; Conservative 0; Mismatches 38; Indels 10; Gaps 4;

QY 513 GCGAGATGTCCTTAACAAGAGCTGATACAGAGATGGAGTTGGTTATGTTAAATGC 572
 DB 2 GCGAGATGTCCTTAACAAGAGCTGATACAGAGATGGAGTTGGTTATGTTAAATGC 61
 QY 573 CAGCTATACCTGTCCAAAGCTGACTTATATCCCGAGTCTCATTTCTGCGTGTGCTGT 632
 DB 62 CAGCTATACCTGTCCAAAGCTGACTTATATCCCGAGTCTCATTTCTGCGTGTGCTGT 121
 QY 633 GAATGACAGCTTTTGTGAAAAATTTTGCCTGTGTGGAACAAATCATATTTTCAATGA 692
 DB 122 GAATGACAGCTTTTGTGAAAAATTTTGCCTGTGTGGAACAAATCATATTTTCAATGA 181
 QY 693 GAAAGCAAAAGCTTCATAGATGTGTTTATGTGACACTGTTTATGCTGGATCTCCGCTC 752
 DB 182 GAAAGCAAAAGCTTCATAGATGTGTTTATGTGACACTGTTTATGCTGGATCTCCGCTC 241
 QY 753 GCGCACATCGCTATCGCTTACATCATGAGAGATGAGACATCTTTTATGATGAAGCTTA 812
 DB 242 GCGCACATCGCTATCGCTTACATCATGAGAGATGAGACATCTTTTATGATGAAGCTTA 301
 QY 813 CAGATTGTGAAAAAGAAAGAAAGCTTATATTTCCAACTTCAATTTTCTGGGCAACT 872
 DB 302 CAGATTGTGAAAAAGAAAGAAAGCTTATATCTCCAACTTCAATTTTCTGGGCAACT 361
 QY 873 CTTGACTATGAGAAAGATTAAAGACAGACTGAGATCAGGGCCAAAGACAACT 932

```

Db      362 CCGGACTATGAGAGAGATTAAGAACGACATCGAGCTCAGGGCCAAAGCAAACT 421
QY      933 CAAGCTGCTGACCTGAGAGAGCCAAATGAACCTGCTCCTGCTCTCAGAGGGTGACA 992
Db      422 CAAGCTGCTGACCTGAGAGAGCCAAATGAACCTGCTCCTGCTCTCAGAGGGTGACA 481
QY      993 GAAAGGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAGAG 1052
Db      482 GAAAGGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 541
QY      1053 ACAAGGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 1112
Db      542 ACAAGGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 601
QY      1113 AGAGGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 1172
Db      602 AGAGGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 661
QY      1173 AGAGGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 1231
Db      662 AGAGGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 721
QY      1232 CCAAGCTGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 1287
Db      722 CCAAGCTGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 781
QY      1288 TACAAACCTTCCACTACTCTGATGAGAGAGCCAAAGC-----TATGCAAGTCTCCCTG 1342
Db      782 TACAAACCTTCCACTACTCTGATGAGAGAGCCAAAGC-----TATGCAAGTCTCCCTG 841
QY      1343 TTGAGGAGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 1402
Db      842 TTGAGGAGAGAGAGCCCTCTGAGTCAACCTGCGGAGTCTGATCTGATCTGATGAGAG 901
QY      1403 C 1403
Db      902 C 902

RESULT 8
LOCUS   BI821804               836 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 60305889F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176724 5',
          mRNA sequence.
ACCESSION BI821804
VERSION   BI821804.1 GI:15933354
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 836)
          NIH-MGC http://mgs.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: c9apbs-r@mail.nih.gov
          Tissue Procurement: Life Technologies, Inc.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Inceye Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: LHAM1440 row: f column: 21
          High quality sequence stop: 805.
          Location/Qualifiers
            1..836
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5176724"

```

```

/1ab host="DH10B"
/clone 11b="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match      34.9%; Score 695.6; DB 12; Length 836;
Best Local Similarity 96.1%; Pred. No. 1,7e-184;
Matches 809; Conservative 0; Mismatches 24; Indels 9; Gaps 9;

QY      147 CTGCTCAAGCTATGAGAGAGAGTTGCAACAGAGCAAAAGTGTTATTACAGAGCTCAT 206
Db      1 CTGCTCAAGCTATGAGAGAGAGTTGCAACAGAGCAAAAGTGTTATTACAGAGCTCAT 60
QY      207 CCAGCATTCAGCGAAACATPAGGTTGACATTTAGTCAGTCAGAGAGTTGTATTAAGA 266
Db      61 CCAGCATTCAGCGAAACATPAGGTTGACATTTAGTCAGTCAGAGAGTTGTATTAAGA 119
QY      267 TCAAGCTCCCAAGATGTTGCTCTCTCTTGAAGAGTGTCTTCTACTGTACTTCTGGG 326
Db      120 TCAAGCTCCCAAGAGT-GTCCCTCTCTCTTGAAGAGTGTCTTCTACTGTACTTCTGGG 178
QY      327 TAACTGAGAGAGAGTTCAACTCTGTTCAACCGTCTGAGAGAGAGGTTGCTGAGTTCTC 386
Db      179 TAACTGAGAGAGAGTTCAACTCTGTTCAACCGTCTGAGAGAGGTTGCTGAGTTCTC 238
QY      387 TCGTGTGTTCCCTGAGCTCTGTGAGAGAAATTCACCTAGTCCCTACCTGATTTCTCA 446
Db      239 TCGTGTGTTCCCTGAGCTCTGTGAGAGAAATTCACCTAGTCCCTACCTGATTTCTCA 298
QY      447 GCTTGTGTTACCTGTTGCGCAACATTTGGCCCAACCCGAAATTTCCCAATTTTATCTGG 506
Db      299 GCTTGTGTTACCTGTTGCGCAACATTTGGCCCAACCCGAAATTTCCCAATTTTATCTGG 358
QY      507 CTGCGACGAGAGATGCTCTCAACAGAGAGCTGATGACAGAGATGGATGGATGGATGGT 566
Db      359 CTGCGACGAGAGATGCTCTCAACAGAGAGCTGATGACAGAGATGGATGGATGGATGGT 418
QY      567 AATGCGAGCTATACCTGTCAAAGGCTGACCTTATCCCGAGTCTCATTTCTGCGGTG 626
Db      419 AATGCGAGAGATACCTGTCAAAGGCTGACCTTATCCCGAGTCTCATTTCTGCGGTG 478
QY      627 GCTGTGATGACAGCTTTTGTGAGAAATTTTGGCGTGTGGACAAATCACTAGATT 686
Db      479 GCTGTGATGACAGCTTTTGTGAGAAATTTTGGCGTGTGGACAAATCACTAGATT 538
QY      687 CATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 746
Db      539 CATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
QY      747 CCGCTCCGACACCATGCTATCGCTTACATCATGAGAGAGAGAGAGAGAGAGAGAGAG 806
Db      598 CCGCTCCG-CACATGCGTATG-CATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 655
QY      807 AGCTTACAGATTTTGAAGAA-AAAAAGCTTATATCTCCAAATTTTCT-G 864
Db      656 AGCTTACAGATTTTGAAGAA-AAAAAGCTTATATCTCCAAATTTTCTG 715
QY      865 GGCACATCTGAGCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 924
Db      716 GGCACATCTGAGCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
QY      925 AGCAAACTCAAGCTGTGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 984
Db      775 AGCTTAATCAAGCTGTGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 833

```

QY 985 GG 986
Db 834 GG 835

RESULT 9

LOCUS

B0770036 862 bp mRNA linear EST 26-JUL-2002
UI-M-F10-Dyt-o-24-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone

DEFINITION

IMAGE:5702255 5', mRNA sequence.
B0770036

ACCESSION

EST 26-JUL-2002
B0770036.1 GI:21978510

VERSION

EST 26-JUL-2002
Mus musculus (house mouse)

KEYWORDS

SOURCE

ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Email: cgabbs-remail.nih.gov
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

source

Seg primer: PYX-5.

Location/Qualifiers

1..862
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5702255"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_11b="NIH_BMAP_F10"
/note="Organ: Brain; Vector: PYX-Asc; Site 1: Ecor I,
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with Not I and then cloned
directionally into PYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCCACGAC. This library was created for the University
Iowa Brain Anatomy Project (BMAP), 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemm Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match 33.1%; Score 660.6; DB 13; Length 862;
Best Local Similarity 86.8%; Pred. No. 1.3e-174;
Matches 752; Conservative 0; Mismatches 102; Indels 12; Gaps 2;

QY 347 ACTCTGTCACCTCTGTCAGGTCGTCAGTTCCTGTTCTTCCCTGACCTCT
Db 9 ACTCTGTCACCTCTGTCAGGTCGTCAGTTCCTGTTCTTCCCTGACCTCT
QY 407 GTGAGGAAATGCACTTATGTCCTTACCTGATTTCTCAGCCTTCTTACCTGTCGA
Db 69 GTGAGGAAATGCACTTATGTCCTTACCTGATTTCTCAGCCTTCTTACCTGTCGA 128

QY 467 ACATTGGGCAACCCGAATTCCTCCCAATCTTTATCTTGGTCGACGAGATGCTCA 526
Db 129 ACATTGGGCAACCCGAATTCCTCCCAATCTTTATCTTGGTCGACGAGATGCTCA 188
QY 527 ACAAGAGCTGATACAGCAAGATGGATGGTATGTTAAATGACGATATACCTGTC 586
Db 189 ACAAGAGCTGATACAGCAAGATGGATGGTATGTTAAATGACGATATACCTGTC 248
QY 587 CAAAGCTGACTTATCCCGAGTCTCATTTCCGCGTGTGCTGTGAATGACAGCTTT 646
Db 249 CAAAGCTGACTTATCCCGAGTCTCATTTCCGCGTGTGCTGTGAATGACAGCTTT 308
QY 647 GTGAGAAATTTTCCCGTGGTGGACAAATCAGTATGTTCAATTGAGAAAGCAAGCT 706
Db 309 GTGAGAAATTTTCCCGTGGTGGACAAATCAGTATGTTCAATTGAGAAAGCAAGCT 368
QY 707 CCAATGATGTTGTCGTCGTCATGTTAGCTGGGATCTCCCGCTCCGACATCGCTA 766
Db 369 CCAATGATGTTGTCGTCGTCATGTTAGCTGGGATCTCCCGCTCCGACATCGCTA 428
QY 767 TCGGCTCATCATGAAAGAGATGACATGCTTTAGATGAGCTTACAGATTTGGAAG 826
Db 429 TCGGCTCATCATGAAAGAGATGACATGCTTTAGATGAGCTTACAGATTTGGAAG 488
QY 827 AAAAAACCTACTATATCTCCAAACTTCAATTTCTGGGCAACTCTGTGACTATGAGA 886
Db 489 AAAAAACCTACTATATCTCCAAACTTCAATTTCTGGGCAACTCTGTGACTATGAGA 548
QY 887 AGAAGATTAAGAACCAAGCTGAGATCAGGACCAAGAGCAAACTGACCTGTCAC 946
Db 549 AGAAGATTAAGAACCAAGCTGAGATCAGGACCAAGAGCAAACTGACCTGTCAC 608
QY 947 TGGAGAGGCAATATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1006
Db 609 TGGAGAGGCAATATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 668
QY 1007 CCTGAGTCCACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1066
Db 669 CCTGAGTCCACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 725
QY 1067 ATCCGCGAGGTCGCCAGGTCGCCAGGTCGCCAGGTCGCCAGGTCGCCAGGTCGCC 1126
Db 726 ATCCGCGAGGTCGCCAGGTCGCCAGGTCGCCAGGTCGCCAGGTCGCCAGGTCGCC 776
QY 1127 TGTGACAGGCGCTCAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1186
Db 777 TGTGACAGGCGCTCAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 836
QY 1187 TCAAGCCTTCCTTCTCTGATATC 1212
Db 837 TCAAGCCTTCCTTCTCTGATATC 862

RESULT 10

LOCUS

BE897795 920 bp mRNA linear EST 20-OCT-2000
60148457F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3923247 5',
mRNA sequence.

DEFINITION

BE897795
BE897795.1 GI:10363618

ACCESSION

EST.
BE897795.1

KEYWORDS

EST.
Homo sapiens (human)

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 920)
Mammalia; Eutheraia; Primates; Catarrhini; Homindae; Homo.

AUTHORS

NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF


```

QY 289 TCTCTCTGAGACTGTTTCTCACTGTAATCTTGGAATACTGAGAAAGACTTCAAC 348
Db 533 TCTCTCTCTGAGACTGTTTCTCACTGTAATCTTGGAATACTGAGAAAGACTTCAAC 474
QY 349 TCTGTTGACCTGCTTGAGAGGTGAGTGTGCTGAGTCTCTGCTGTTTCTCCCTGCTCTGT 408
Db 473 TCTGTTGACCTGCTTGAGAGGTGAGTGTGCTGAGTCTCTGCTGTTTCTCCCTGCTCTGT 414
QY 409 GAGAGAAATATCACTGTAATCTTCACTGTAATCTTGGAATACTGAGAAAGACTTCAAC 468
Db 413 GAGAGAAATATCACTGTAATCTTCACTGTAATCTTGGAATACTGAGAAAGACTTCAAC 354
QY 469 ATTGGGCAACCCGGAATCTTCCCAATCTTGAATCTTGGAATACTGAGAAAGACTTCAAC 528
Db 353 ATTGGGCAACCCGGAATCTTCCCAATCTTGAATCTTGGAATACTGAGAAAGACTTCAAC 294
QY 529 AAGAGCTGATACAGAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 588
Db 293 AAGAGCTGATACAGAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 234
QY 589 AAGCTGACTTATATCCCGAGTCTCATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
Db 233 AAGCTGACTTATATCCCGAGTCTCATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 174
QY 649 GAGAAATATTTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
Db 173 GAGAAATATTTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 114
QY 709 AATGATGTTTCTAGTCACTGTTAGCTGGAATCTCCGCTCCGCTCCGCTCCGCTCCGCT 768
Db 113 AATGATGTTTCTAGTCACTGTTAGCTGGAATCTCCGCTCCGCTCCGCTCCGCTCCGCT 54
QY 769 GCTTACATCAAGAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 821
Db 53 GCTTACATCAAGAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1

```

```

RESULT 12
LOCUS BX479029 663 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686J22208.r1.686 (synonym: hicc3) Homo sapiens cDNA clone
ACCESSION DKFZp686J22208.5, mRNA sequence.
VERSION BX479029
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
JOURNAL EST (Bahr, A., Lauber, J., Mewes, H.W., Weil, B., et al.)
COMMENT Unpublished (2003)
CONTACT: MIPS
MIPS
Ingolstaedter Landstr. 1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de,
consortium of the German Genome Project.
No 5' sequence available.
This clone (DKFZp686J22208) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY, Email: clone@rzpd.de.

```

FEATURES

```

source
1..663
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686J22208"

```

```

/dev_stage="adult"
/lab_host="DH10B"
/clone_11b="686 (synonym: hicc3)"
/note="Vector: pTIPLEX2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

```

```

Query Match 31.7%; Score 632; DB 13; Length 663;
Best Local Similarity 98.6%; Pred. No. 1.3e-16;
Matches 647; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

```

```

QY 665 GGTGAGCAATATCATGATTTTCAAT--GAGAAAGCAAAAGCTCCATGATGATGAT 722
Db 2 GGTGAGCAATATCATGATTTTCAATGAGCAAAAGCAAAAGCTCCATGAGAGCTTCT 61
QY 723 AGTGCATGTTTACCTGGAATCTCCGCTCCGCAACATGCTATGCTTATGCTATCATGAA 782
Db 62 AGTGCATGTTTACCTGGAATCTCCGCTCCGCAACATGCTATGCTTATGCTATCATGAA 121
QY 783 GAGATGAGCAATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
Db 122 GAGATGAGCAATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181
QY 843 ATCTCAAACTTCAATTTCTGGGCCCACTCTGCACTATGAGAGAAAGATTAGAACCA 902
Db 182 ATCTCAAACTTCAATTTCTGGGCCCACTCTGCACTATGAGAGAAAGATTAGAACCA 241
QY 903 GACTGAGCAATGAGGAGCAAAAGCAAACTCAAGCTGCTGCACTGAGAGCAAAATGA 962
Db 242 GACTGAGCAATGAGGAGCAAAAGCAAACTCAAGCTGCTGCACTGAGAGCAAAATGA 301
QY 963 ACTGTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
Db 302 ACTGTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
QY 1023 TGGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
Db 362 TGGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
QY 1083 CAGCTGCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
Db 422 CAGCTGCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
QY 1143 TGGGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
Db 482 TGGGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
QY 1203 TCTGATATCAATATCAATATCAATATCAATATCAATATCAATATCAATATCAATATCA 1262
Db 542 TCTGATATCAATATCAATATCAATATCAATATCAATATCAATATCAATATCAATATCA 601
QY 1263 CTCATCAGAGATGCTTGGAAATACACAACTTCCACTCTGATGAGGACCA 1318
Db 602 CTCATCAGAGATGCTTGGAAATACACAACTTCCACTCTGATGAGGACCA 657

```

```

RESULT 13
LOCUS B0670989 891 bp mRNA linear EST 15-JUN-2002
DEFINITION AGENCOURT 8039377 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6211589
ACCESSION B0670989
VERSION B0670989.1 GI:21781823
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://nigc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

```


Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: ARCC
 CDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM374 Row: 6 Column: 06
 High quality sequence stop: 633.
 Location/Qualifiers
 1. .891
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6211589"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH MGC 102"
 /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; CDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-CDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies). Note: this is a NIH_MGC Library."

FEATURES

source

Query Match 31.4%; Score 627; DB 13; Length 891;
 Best Local Similarity 100.0%; Pred. No. 4,1e-165;
 Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

1369 GAAACCAAGTCTGTATAGAGAGAGCCAGCATCCCAAGAGCTGCAGACCGGCGCTT 1428
 1 GAAACCAAGTCTGTATAGAGAGAGCCAGCATCCCAAGAGCTGCAGACCGGCGCTT 60
 1429 TCAGACAGCCAGAGCAAGCATTCGCTGCAAGAACAGAGCAAGTGGCAACCGCCAG 1488
 61 TCAGACAGCCAGAGCAAGCATTCGCTGCAAGAACAGAGCAAGTGGCAACCGCCAG 120
 1489 AGGTCCCTTTATCTCCATGCAATCGAAGTGGAGCGTGGAGGCAATTACACACAGC 1548
 121 AGGTCCCTTTATCTCCATGCAATCGAAGTGGAGCGTGGAGGCAATTACACACAGC 180
 1549 TTCTTTTGGGCTTTTCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1608
 181 TTCTTTTGGGCTTTTCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 240
 1609 AAGGCTGGCACTCGATATCTTGGCCCCCAAGCTCTACCCCTTCCTGCAGCAGCAGC 1668
 241 AAGGCTGGCACTCGATATCTTGGCCCCCAAGCTCTACCCCTTCCTGCAGCAGCAGC 300
 1669 TGGTATTTTSCACAGAGCTCTCACTTCTACCTGCGCTCAGCAATCTACGAGGAGT 1728
 301 TGGTATTTTSCACAGAGCTCTCACTTCTACCTGCGCTCAGCAATCTACGAGGAGT 360
 1729 GCCAGTTACTCTGCTCAAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 1788
 361 GCCAGTTACTCTGCTCAAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 420
 1789 GTGCGCAGGCGGAGAGCAAGTGAAGAGCTGAGCTGGGAGAGCTGGCATGAAGAG 1848
 421 GTGCGCAGGCGGAGAGCAAGTGAAGAGCTGAGCTGGGAGAGCTGGCATGAAGAG 480
 1849 AGCCCTTTTGAAGAGCTTAAACGAGAGCTGCCAATGGAATTTGGAGAGAGATC 1908
 481 AGCCCTTTTGAAGAGCTTAAACGAGAGCTGCCAATGGAATTTGGAGAGAGATC 540
 1909 ATGTCAAGAGAGCTGACGAGAGAGCTGGGAAAGTGGGAGTCAAGTCAAGCTTTTCG 1968
 541 ATGTCAAGAGAGCTGACGAGAGAGCTGGGAAAGTGGGAGTCAAGTCAAGCTTTTCG 600

QY 1669 GGCAGCATGAATCATTCAGGCTTCC 1995
 DB 601 GGCAGCATGAATCATTCAGGCTTCC 627

RESULT 14
 BU704078
 LOCUS
 DEFINITION
 BU704078 769 bp mRNA linear EST 15-JUL-2003
 UI-M-FOO-bar-n-23-0-UI.r1 NIH_BMAP_FOO Mus musculus cDNA clone
 IMAGE:6406486 5', mRNA sequence.
 BU704078
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 769)
 NIH-MGC http://img.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Seq primer: pyx-5.
 Location/Qualifiers
 1. .769
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6406486"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1ib="NIH BMAP_FOO"
 /note="Organ: Brain; Vector: pyx-Asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction. Ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pyx-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is TGAAGAGGC. This library was created for the University
 Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemm Chin, Ph.D.,
 program coordinator."

FEATURES

source

Query Match 31.2%; Score 623.2; DB 13; Length 769;
 Best Local Similarity 88.8%; Pred. No. 4,4e-164;
 Matches 673; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

ORIGIN

274 TCCCAAGATGTCCTCTCTCTCAGACTGTTTCTCACTGTAATCTTGGGTAAGTCC 333
 3 TCCCAAGATGTCCTCTCTCTCAGACTGTTTCTCACTGTAATCTTGGGTAAGTCC 62
 334 GAGAGAGCTTCAACTCTGTCACCTGTCAGAGTGGGTTGCTGAGTCTCTCGTTGT 393
 63 GAGAGAGCTTCAACTCTGTCACCTGTCAGAGTGGGTTGCTGAGTCTCTCGTTGT 122

QY 394 TTCCTGGGCTCTGTGAAGAAATCACTAGTCCCTAGCTGATTTCTACGCTTGC 453
DB 123 TTCCTGGGCTCTGTGAAGAAATCACTAGTCCCTAGCTGATTTCTACGCTTGC 182
QY 454 TTACTCTGTGGCCAAATTTGGCCAAATTTCTTCCCAATCTTTAATCTTGGCTGCAG 513
DB 183 TTACTCTGTGGCCAAATTTGGCCAAATTTCTTCCCAATCTTTAATCTTGGCTGCAG 242
QY 514 CGAGATGCTCTCAACAAAGAGCTGATACAGCAATGGAGATTTGGTTATGTTAATGCC 573
DB 243 CGAGATGCTCTCAACAAAGAGCTGATACAGCAATGGAGATTTGGTTATGTTAATGCC 302
QY 574 AGCTATACCTGTCCAAAGCCTGACTTTATCCCGAGTCTCAATTCCTGCGGTGCTGCTG 633
DB 303 AGCAATACCTGTCCAAAGCCTGACTTTATCCCGAGTCTCAATTCCTGCGGTGCTGCTG 362
QY 634 AATGACAGCTTTGTGAGAAATTTTGGCGGTGTTGACAAATGATGATTTCAATTGAG 693
DB 363 AATGACAGCTTTGTGAGAAATTTTGGCGGTGTTGACAAATGATGATTTCAATTGAG 422
QY 694 AAAGCAAAAGCCTCCAAATGATGATGTTCTGATGCACTGTTAGCTGGATTTCCGCTGC 753
DB 423 AAAGCAAAAGCCTCCAAATGATGATGTTCTGATGCACTGTTAGCTGGATTTCCGCTGC 482
QY 754 GCCACATTCGCTATTCGCTATCATGATGAGAGATGACATGTTTATGATGAGCTTAC 813
DB 483 GCCACATTCGCTATTCGCTATCATGATGAGAGATGACATGTTTATGATGAGCTTAC 542
QY 814 AGATTTGTGAAGAAATTTGACCTACTATCTTCCAAATTTTGTGGCCAACTC 873
DB 543 AGATTTGTGAAGAAATTTGACCTACTATCTTCCAAATTTTGTGGCCAACTC 602
QY 874 CTGACATGATGAGAAATTTGATTAAGAACAGATGAGCATCAGGGCCAAAGCAACTC 933
DB 603 ATGACATGATGAGAAATTTGATTAAGAACAGATGAGCATCAGGGCCAAAGCAACTC 662
QY 934 AAGCTGTGACCTGTGAGAAAGCCAAATGAACTGTCCTGCTGCTGACAGGCTGACAG 993
DB 663 AAGCTGTGACCTGTGAGAAAGCCAAATGAACTGTCCTGCTGCTGACAGGCTGACAG 722
QY 994 AAAAGGAGAGAGCCCTCTGATTCACCTGTGCGCACTC 1031
DB 723 AAGAGTGAAGCTGTCTCTGATCCACTCTGTGCAACTC 760

RESULT 15
BI917706/c 660 bp mRNA linear EST 16-OCT-2001
LOCUS 60318366F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5247884 5',
DEFINITION mRNA sequence.
ACCESSION BI917706
VERSION BI917706.1 GI:16181504
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: L1AM1625 row: k column: 21
High quality sequence stop: 660.
Location/Qualifiers

source 1..660
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5247884"
/lab_host="DH10B"
/clone_11b="NIH MGC 121"
/note="Organ: brain; Vector: pCMV-Sport6; Site: 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

Query Match 30.8%; Score 615.4; DB 12; Length 660;
Best Local Similarity 98.2%; Pred. No. 6.3e-162;
Matches 657; Conservative 0; Mismatches 1; Indels 11; Gaps 3;

QY 707 CCATGATGATGTTCTAGTGGACACTGTTAGCTGGATCTCCGCTGC-GCCACATCGCT 765
DB 660 CCATGATGATGTTCTAGTGGACACTGTTAGCTGGATCTCCGCTGCAGATCGCT 601
QY 766 ATGCGCTACATCATGAGAGATGACATGCTTTTATGATGAGCTTACAGATTTGTGAA 825
DB 600 ATGCGCTACATCATGAGAGATGACATGCTTTTATGATGAGCTTACAGATTTGTGAA 541
QY 826 GAAAAAAGCCTATCTATCTCCA-AACTTCAATTTTGTGGCCAACTCTGACTATGA 884
DB 540 GAAAAAAGCCTATCTATCTCCA-AACTTCAATTTTGTGGCCAACTCTGACTATGA 481
QY 885 GAAAGATTTAAGAACAGACTGAGCATCAGGGCCAAAGCAACTCAAGCTGTGCA 944
DB 480 GAAAGATTTAAGAACAGACTGAGCATCAGGGCCAAAGCAACTCAAGCTGTGCA 421
QY 945 CCTGGAAGGCCAATGAACTGTCTCTGCTGTCTCAGAGGTGAGACAGAAACAGAGAC 1004
DB 420 CCTGGAAGGCCAATGAACTGTCTCTGCTGTCTCAGAGGTGAGACAGAAACAGAGAC 361
QY 1005 GCCCCTCAGTCCACCTGTGCGCACTGTGCTACCTCAGAGGACAGAGCAAAAGCCCGT 1064
DB 360 GCCCCTCAGTCCACCTGTGCGCACTGTGCTACCTCAGAGGACAGAGCAAAAGCCCGT 301
QY 1065 GCATCCCCGACGCTGCCAGCGTGCAGGCTGACGCTGTGTTAGAGACAGCC 1124
DB 300 GCATCCCCGACGCTGCCAGCGTGCAGGCTGACGCTGTGTTAGAGACAGCC 250
QY 1125 GCTGTGATCAGGCGCTCAGTGGCTGACCTGTGCGGACAGAGGCTGGAACAGCAATTA 1184
DB 249 GCTGTGATCAGGCGCTCAGTGGCTGACCTGTGCGGACAGAGGCTGGAACAGCAATTA 190
QY 1185 GCTCAGGCTTCTTCTCTGATATCAATCAATGTTTCAATTCAGCCAGATGGCAGC 1244
DB 189 GCTCAGGCTTCTTCTCTGATATCAATCAATGTTTCAATTCAGCCAGATGGCAGC 130
QY 1245 ATCTTACATGCTTCTCCATCAGAAAGATGCTTGAATATCTAACAACCTTCCACTAC 1304
DB 129 ATCTTACATGCTTCTCCATCAGAAAGATGCTTGAATATCTAACAACCTTCCACTAC 70
QY 1305 TCTGATGAGACCAACAGCTATGCGATTCCTCCCTGTTCAGAGACTATCGAGAGAG 1364
DB 69 TCTGATGAGACCAACAGCTATGCGATTCCTCCCTGTTCAGAGACTATCGAGAGAG 10
QY 1365 TCCGGAAC 1373
DB 9 TCCGGAAC 1

Search completed: June 22, 2004, 02:51:54

FEATURES

Tue Jun 22 16:59:31 2004

us-10-029-345a-108_copy_538_2532.rst

Page 13

Job time : 7901.16 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 12:17:10 ; Search time 1159.45 Seconds
(without alignments)
7309.622 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccatgagatgatg9.....tggaatcatgagctcc 1995

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002s:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1995	100.0	5450	6	ACC60559
2	1995	100.0	5450	6	ACC60559
3	1991.8	99.8	3059	6	AA515768
4	1991.8	99.8	3496	6	ABK47596
5	1991.8	99.8	3544	5	AA514639
6	1991.8	99.8	3544	6	ABK49402
7	1991.8	99.8	3766	6	ABK14474
8	1991.8	99.8	4790	6	ABN83966
9	1991.8	99.8	5145	5	ABV20833
10	1991.8	99.8	5145	5	ABV21080
11	1991.8	99.8	5145	5	ABV26680
12	1991.8	99.8	5145	5	ABV20978
13	1991.8	99.8	5145	5	ABV21092
14	1991.8	99.8	5145	5	ABV21312
15	1991.8	99.8	5145	5	ABV21316
16	1991.8	99.8	5145	5	ABV26826
17	1991.8	99.8	5145	5	ABV27131
18	1991.8	99.8	5145	5	ABV26923
19	1991.8	99.8	5145	5	ABV27135
20	1991.8	99.8	5145	5	ABV28657
21	1991.8	99.8	5145	5	ABV22827
22	1991.8	99.8	5145	5	ABV26934
23	1990.2	99.8	2732	4	AA09492

24	1990.2	99.8	3104	6	ABN59704	Abn59704 Novel hum
25	1988.6	99.7	2102	7	ADA53105	Ada53105 Human cod
26	1988.6	99.7	2966	4	AAH99685	Aah99685 Human pro
27	1988.6	99.7	5111	6	ACC60521	Acc60521 polynucle
28	1968.8	98.7	2118	4	AAE30479	AAE30479 Human pro
29	1653.8	82.9	3332	6	ABK48378	Abk48378 CDNA enco
30	1554.8	77.9	2756	6	ACC60560	Acc60560 Polynucle
31	727.6	36.5	749	4	AAH06539	Aah06539 Human CDN
32	417.2	20.9	425	5	ABV10726	Abv10726 Human pro
33	415.2	20.8	467	5	ABV40958	Abv40958 Human pro
34	415.2	20.8	467	5	ABV31891	Abv31891 Human pro
35	415.2	20.8	467	5	ABV40849	Abv40849 Human pro
36	411.2	20.6	418	5	ABV11252	Abv11252 Human pro
37	410.2	20.6	461	5	ABV40981	Abv40981 Human pro
38	410.2	20.6	461	5	ABV23398	Abv23398 Human pro
39	410.2	20.6	461	5	ABV41327	Abv41327 Human pro
40	401.2	20.1	408	5	ABV10907	Abv10907 Human pro
41	399.2	20.0	438	5	ABV32059	Abv32059 Human pro
42	395.4	19.8	427	5	ABV10887	Abv10887 Human pro
43	387.8	19.4	2377	7	ACA64956	Aca64956 Human pro
44	387.8	19.4	2377	7	ABX10760	Abx10760 Human pro
45	382	19.1	2453	2	AAT86758	Aat86758 CDNA of t

ALIGNMENTS

RESULT 1	ACCG60559	standard; cDNA; 5450 BP.
AC	ACCG60559;	
DT	19-JUN-2003	(first entry)
XX	Polynucleotide relating to the invention SEQ ID NO: 108.	
DE	Gene; sei; antiproliferative; hepatotropic; nephrotropic; antiarthritic;	
XX	antiproliferative; cardiac; cytosolic; gene therapy; liver disease;	
KW	proliferative disorder; renal failure; cardiovascular disorder;	
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;	
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	WO200257460-A2.	
PN	25-JUL-2002.	
PD	20-DEC-2001; 2001WO-US050459.	
XX	20-DEC-2001; 2000US-025668P.	
XX	31-MAR-2001; 2001US-028018P.	
PR	01-MAY-2001; 2001US-0287735P.	
PR	05-JUN-2001; 2001US-029548P.	
PR	25-JUN-2001; 2001US-0300465P.	
XX	(BRIM) BRISTOL-MYERS SQUIBB CO.	
PA	Jackson DG, Feder J, Nelson T, Minter G, Ramanathan C, Lee L;	
XX	Stemerson N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;	
PI	Krysstek S, Mcatee P, Suchard S, Barnes D;	
XX	WPI; 2002-599721/64.	
DR	P-PSDB; ABR52381.	
XX	Novel polynucleotides encoding human phosphatase polypeptides useful in	
PT	the prevention or treatment of e.g. proliferative and cardiovascular	
PT	disorders.	
XX	Claim 1; Fig 13; 801p; English.	
PS	The invention relates to a novel isolated nucleic acid comprising a	
XX		
CC		

CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic, CC
CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC ameliorating the invention is useful for preventing, treating or
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX
SQ Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;

Query Match 100.0%; Score 1995; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCATGACATGATTTGAACTCAATTTGTTACTGAGAGGTGGTCTGCTGAA 60
DB 538 ATGGCCCATGAGATGATTTGAACTCAATTTGTTACTGAGAGGTGGTCTGCTGAA 60
QY 61 AGTGAACGGAAAAAGTCTGCTAATGATAGCCGGCATTTGTGAATACATATCC 120
DB 598 AGTGAACGGAAAAAGTCTGCTAATGATAGCCGGCATTTGTGAATACATATCC 120
QY 121 CACATTTTGGAAAGCATTAATATCACTGCTCAAGCTTATGAGCCAAAGTTGCAACG 180
DB 658 CACATTTTGGAAAGCATTAATATCACTGCTCAAGCTTATGAGCCAAAGTTGCAACG 180
QY 181 GACAAAGTGTAAATACAGAGCTCATCAGCATTCAGCCGAAACATAGTTGACATTGAT 240
DB 718 GACAAAGTGTAAATACAGAGCTCATCAGCATTCAGCCGAAACATAGTTGACATTGAT 240
QY 241 TGCAGTCAGAAAGTTGATGATTAACATCAAGCTCCCAAGATGTTCTCTCTCTCA 300
DB 778 TGCAGTCAGAAAGTTGATGATTAACATCAAGCTCCCAAGATGTTCTCTCTCTCA 300
QY 301 GACTGTTTCTCACTGACTCTTGGGGTAAACGAGAAAGCTTCAACTCTGTTCACTG 360
DB 838 GACTGTTTCTCACTGACTCTTGGGGTAAACGAGAAAGCTTCAACTCTGTTCACTG 360
QY 361 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGAAATCC 420
DB 898 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGAAATCC 420
QY 421 ACTCTAGTCCCTACCTGCAATTTCTAGCTTGTCTGCTGCTGCTGCTGCTGCTGCT 480
DB 958 ACTCTAGTCCCTACCTGCAATTTCTAGCTTGTCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CGAATTTCTCCCAATTTTATCTTGGCTGCGACGAGATGTCCTCAACAAGAGCTGATA 540
DB 1018 CGAATTTCTCCCAATTTTATCTTGGCTGCGACGAGATGTCCTCAACAAGAGCTGATA 540
QY 541 CAGCAAGATGAGATGTTATGTTAAATGCGCTATACCTGTTCCAAAGCTGACTT 600
DB 1078 CAGCAAGATGAGATGTTATGTTAAATGCGCTATACCTGTTCCAAAGCTGACTT 600
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 1138 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 CCGTGTGTGACAATAGTATGATTTGATGAGAAAGCAAAAGCTTCCAAAGCTGATGTT 720
DB 1198 CCGTGTGTGACAATAGTATGATTTGATGAGAAAGCAAAAGCTTCCAAAGCTGATGTT 720
QY 721 CTATGTGACTGTTTAAAGTGGATCTCCGCTCCGACCATGCTATGCGCTCAATCATG 780
DB 1258 CTATGTGACTGTTTAAAGTGGATCTCCGCTCCGACCATGCTATGCGCTCAATCATG 780
QY 781 AAGAGATGACATGCTTTTAAATGATGAAGCTTACAGATTTGTGAAAGAAAAAGCTTACT 840

DB 1318 AAGAGATGACATGCTTTTAAATGATGAAGCTTAAAGATTTGTGAAAGAAAAAGCTTACT 1377
QY 841 ATATCTCAAACTCAATTTTCTGGGCCCACTCTCTGACATTAAGAAAGATTTAAGAAC 900
DB 1378 ATATCTCAAACTCAATTTTCTGGGCCCACTCTCTGACATTAAGAAAGATTTAAGAAC 900
QY 901 CAGACTGAGATCAAGGCGCAAGAAAGCAAACTCAAGCTGTGCACTGAGAAAGCCAAAT 960
DB 1438 CAGACTGAGATCAAGGCGCAAGAAAGCAAACTCAAGCTGTGCACTGAGAAAGCCAAAT 960
QY 961 GAACCTGTCTCCGTGCTGCTGAGAGGTGACAAAAAGCAAGCGCCCTCACTGCAACC 1020
DB 1498 GAACCTGTCTCCGTGCTGCTGAGAGGTGACAAAAAGCAAGCGCCCTCACTGCAACC 1020
QY 1021 TGTGCGCATCTGCTTACCTGAGAGGCAAGCAAAAGCGCGTGTATCCGCGAGCGCG 1080
DB 1558 TGTGCGCATCTGCTTACCTGAGAGGCAAGCAAAAGCGCGTGTATCCGCGAGCGCG 1080
QY 1081 CCGAGCGTCCAGCGTGTGAGAGGCGCGTGTATGAGAGCAAGCGCGTGTATGAGAGCG 1140
DB 1618 CCGAGCGTCCAGCGTGTGAGAGGCGCGTGTATGAGAGCAAGCGCGTGTATGAGAGCG 1140
QY 1141 AGTGGGTGCACTCTGTCCGAGACAGGCTGAGAAAGCAAGATTAAGTCAAGGCTTCTC 1200
DB 1678 AGTGGGTGCACTCTGTCCGAGACAGGCTGAGAAAGCAAGATTAAGTCAAGGCTTCTC 1200
QY 1201 TCTCTGATATCAATCACTTCAATTAATGAGAGCAAGCAAGCAAGCTTCAATGAGCTTC 1260
DB 1738 TCTCTGATATCAATCACTTCAATTAATGAGAGCAAGCAAGCAAGCTTCAATGAGCTTC 1260
QY 1261 TCTCTGATATCAATCACTTCAATTAATGAGAGCAAGCAAGCTTCAATGAGCTTC 1320
DB 1798 TCTCTGATATCAATCACTTCAATTAATGAGAGCAAGCAAGCTTCAATGAGCTTC 1320
QY 1321 AAGCTATGCAAGTCTTCCCTGTTTCAAGAACTATGAGAGCAAGCTTCCGAAACAGTCT 1380
DB 1858 AAGCTATGCAAGTCTTCCCTGTTTCAAGAACTATGAGAGCAAGCTTCCGAAACAGTCT 1380
QY 1381 GATTAAGAGAAAGCCAGATCCCAAGAGCTGAGACCCGAGGCTTCAAGAGCAAG 1440
DB 1918 GATTAAGAGAAAGCCAGATCCCAAGAGCTGAGACCCGAGGCTTCAAGAGCAAG 1440
QY 1441 AGCAAGCATTTGATGCTGAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1500
DB 1978 AGCAAGCATTTGATGCTGAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1500
QY 1501 TCTCACTGATGAGAGTGGAGCGTGGAGCAATTAACAAGCAAGCTTCTTTTGGG 1560
DB 2038 TCTCACTGATGAGAGTGGAGCGTGGAGCAATTAACAAGCAAGCTTCTTTTGGG 1560
QY 1561 CTTTCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
DB 2098 CTTTCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
QY 1621 TCGGATATCTTGGCCCCCAAGCTTCACTTCCCTGACCAAGCAGCTGATTTTGGC 2157
DB 2158 TCGGATATCTTGGCCCCCAAGCTTCACTTCCCTGACCAAGCAGCTGATTTTGGC 2157
QY 1681 ACAGAGTCTCACTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1740
DB 2218 ACAGAGTCTCACTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1740
QY 1741 GCTTACAGCTGACAGCAGCTGCGCACTTCCGAGAGCAAGCTTATGAGAGCAAGCTT 2277
DB 2278 GCTTACAGCTGACAGCAGCTGCGCACTTCCGAGAGCAAGCTTATGAGAGCAAGCTT 2277
QY 1801 CAGAAAGCAAGTGAAGAGCTGATCCGCGGAGAGCTGAGCTGAGAGAGAGAGAGAGAG 2337
DB 2338 CAGAAAGCAAGTGAAGAGCTGATCCGCGGAGAGCTGAGCTGAGAGAGAGAGAGAGAG 2337
QY 1861 AAGCATTTAAAGCAGAGCTGCAATGAAATTTGAGAGAGCAATGATGCAAGAAC 1920

Db 2398 AGCAGTTTAAAGCAGAGAGCTCCAAATGGAATTTGGAGAGAGCATCATGTCAGAGAAC 2457
 Qy 1921 AGGTCAAGGGAAGCTGGGGAAAGTGGCAATGCTAGCTTTTGGGAGAGTGA 1980
 Db 2458 AGGTCAAGGGAAGCTGGGGAAAGTGGCAATGCTAGCTTTTGGGAGAGTGA 2517
 Qy 1981 ATCATTGAGGTCTCC 1995
 Db 2518 ATCATTGAGGTCTCC 2532

RESULT 2
 ACC60572
 ID ACC60572 standard; cDNA; 5450 BP.
 XX ACC60572;
 XX 19-JUN-2003 (first entry)
 XX
 DE Polynucleotide relating to the invention SEQ ID NO: 147.
 XX
 XX Gene; ss; antiproliferative; hepatocarcinoma; nephrotropic; antiarthritic;
 XX antiproliferative; cardiac; cytosolic; gene therapy; liver disease;
 XX proliferative disorder; renal failure; cardiovascular disorder;
 XX immunological disorder; arthritis; psoriasis; congenital heart defect;
 XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
 XX Homo sapiens.
 XX
 XX MO200257460-A2.
 XX
 XX 25-JUL-2002.
 XX
 XX 20-DEC-2001; 2001MO-US050459.
 XX
 XX 20-DEC-2001; 2000US-0256868P.
 XX 30-MAR-2001; 2001US-0280186P.
 XX 01-MAY-2001; 2001US-0287735P.
 XX 05-JUN-2001; 2001US-0295848P.
 XX 25-JUN-2001; 2001US-0300465P.
 XX
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
 XX Stiemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
 XX Kytelex S, Mcatee P, Suchard S, Banas D;
 XX
 XX WPI; 2002-599721/64.
 XX P-PsDB; ABR52407.
 XX
 XX Novel polynucleotides encoding human phosphatase polypeptides useful in
 XX the prevention or treatment of e.g. proliferative and cardiovascular
 XX disorders.
 XX
 XX Example 7; Fig 19; 801pp; English.

Query Match 100.0%; Score 1995; DB 6; Length 5450;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATGATGGAATCTCAATTTGTTACTGAGAGTGTGGCTCTGCGGAA 60
 Db 538 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACTGAGAGTGTGGCTCTGCGGAA 597
 Qy 61 AGTGAACCGGAAAAGTGTCTGCTTAATGATAGCGGCGCATTTGTGAATCAATACATCC 120
 Db 598 AGTGAACCGGAAAAGTGTCTGCTTAATGATAGCGGCGCATTTGTGAATCAATACATCC 657
 Qy 121 CACATTTTGAAGCCATTAATATCACTGCTCCAGCTTATGAAGGAGTTGCAACAG 180
 Db 658 CACATTTTGAAGCCATTAATATCACTGCTCCAGCTTATGAAGGAGTTGCAACAG 717
 Qy 181 GACAAAGTTTAAATTAAGAGCTCATCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
 Db 718 GACAAAGTTTAAATTAAGAGCTCATCAGCATTCAGCGAAACATTAAGTTGACATTGAT 777
 Qy 241 TGCAGTCAGAGGTTGATGATTAATGATCAAGTCCCAAGATGTGCTCTCTTCA 300
 Db 778 TGCAGTCAGAGGTTGATGATTAATGATCAAGTCCCAAGATGTGCTCTCTTCA 837
 Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAACAGAGAGCTTCAACTGTTCACTG 360
 Db 838 GACTGTTTCTCACTGATCTTCTGGGTAACAGAGAGCTTCAACTGTTCACTG 897
 Qy 361 CTTCAGAGTGGGTTTCTGAGTCTCTGTTGTTTCTGCGCTCTGTGAAGGAAATCC 420
 Db 898 CTTCAGAGTGGGTTTCTGAGTCTCTGTTGTTTCTGCGCTCTGTGAAGGAAATCC 957
 Qy 421 ACTGTAGTCCCTTAACCTGATTTCTGAGCTTGTCTTACCTGTGCAACATTTGGCCCAAC 480
 Db 958 ACTGTAGTCCCTTAACCTGATTTCTGAGCTTGTCTTACCTGTGCAACATTTGGCCCAAC 1017
 Qy 481 CGAATCTTCCCAATCTTATCTTGGCTGCGAGAGATGCTTCAACAGAGCTGATA 540
 Db 1018 CGAATCTTCCCAATCTTATCTTGGCTGCGAGAGATGCTTCAACAGAGCTGATA 1077
 Qy 541 CAGCAGAAATGGAGTTGATTAATGATTAATGAGAGTATACCTGTCCAAAGCTGACTTT 600
 Db 1078 CAGCAGAAATGGAGTTGATTAATGATTAATGAGAGTATACCTGTCCAAAGCTGACTTT 1137
 Qy 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGTGAATGACAGCTTTTGTGAATAATTTG 660
 Db 1138 ATCCCGAGTCTCATTTCTGCTGCTGCTGTGAATGACAGCTTTTGTGAATAATTTG 1197
 Qy 661 CCGTGTGGAGCAATCAGAGATTTCAATGAGAAAGCAAAAGCTCCAAATGAGTGTG 720
 Db 1198 CCGTGTGGAGCAATCAGAGATTTCAATGAGAAAGCAAAAGCTCCAAATGAGTGTG 1257
 Qy 721 CTATGCACTGTTTAACTGGATCTCCGCTCCGACCATTCGCTATCGCTATCATCATG 780
 Db 1258 CTATGCACTGTTTAACTGGATCTCCGCTCCGACCATTCGCTATCGCTATCATCATG 1317
 Qy 781 AAGAGATGAGCATGTTCTTTAGTGAAGCTTACAGATTTGTGAAGAAAAGCACTACT 840
 Db 1318 AAGAGATGAGCATGTTCTTTAGTGAAGCTTACAGATTTGTGAAGAAAAGCACTACT 1377
 Qy 841 ATATCTCAAACTCAATTTTCTGGGCAACCTCTGAGATATGAGAAAGATTAAGAAC 900
 Db 1378 ATATCTCAAACTCAATTTTCTGGGCAACCTCTGAGATATGAGAAAGATTAAGAAC 1437
 Qy 901 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTGTGAGAGCCAAAT 960
 Db 1438 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTGTGAGAGCCAAAT 1497
 Qy 961 GAACCTGTCTCTGCTCTCAAGAGGTGACAGAAAAGGAGAGCGCCCTCACTCAACC 1020
 Db 1498 GAACCTGTCTCTGCTCTCAAGAGGTGACAGAAAAGGAGAGCGCCCTCACTCAACC 1557
 Qy 1021 TGTGCCGACTCTGCTCTCAGAGGAGCAGCAAAAGCGCCGTGATCCCGCAGCGTG 1080

Db	1558	TTGTGCGACTCTGCTTACTCTCAGAGGACAGACAAAGGCGCCGTGCATCCCGCAGCGG	1617
QY	1081	CCGACGCTGCCAGCGTGCAGCGCCCTGCCTGTTAGAGACAGCGCCGCTGTATAGCGCTC	1140
Db	1618	CCGACGCTGCCAGCGTGCAGCGCCGCTGTGTAGAGACAGCGCCGCTGTATAGCGCTC	1677
QY	1141	AGTGGGCTGCACCTGTCCGACGACAGGCTGTGAAGACAGCAATAGCTCAAGGCTCTTC	1200
Db	1678	AGTGGGCTGCACCTGTCCGACGACAGGCTGTGAAGACAGCAATAGCTCAAGGCTCTTC	1737
QY	1201	TCCTGGAATCAATCAATGATTGTAATATTCGACGAGCATGGGAGCATCTTACATGGCTTC	1280
Db	1738	TCCTGGAATCAATCAATGATTGTAATATTCGACGAGCATGGGAGCATCTTACATGGCTTC	1797
QY	1281	TCCTCATCAAGAAATGCTTTTGGAAATCTAACAACTTCCACTACCTGTGATGGAGCAAC	1320
Db	1798	TCCTCATCAAGAAATGCTTTTGGAAATCTAACAACTTCCACTACCTGTGATGGAGCAAC	1857
QY	1321	AAGCTATGCGAGTTCTCCCGCTGTTCAGGAACTATGGAGCAGACTCCGAAACCAATGCT	1380
Db	1858	AAGCTATGCGAGTTCTCCCGCTGTTCAGGAACTATGGAGCAGACTCCGAAACCAATGCT	1917
QY	1381	GATPAGAGGAAAGCCAGCATCCCAAGAAAGCTGACAGCCGCGAGCGCTTCAGACAGCGAG	1440
Db	1918	GATPAGAGGAAAGCCAGCATCCCAAGAAAGCTGACAGCCGCGAGCGCTTCAGACAGCGAG	1977
QY	1441	AGCAAGGAAATGATTCGGTTCAGAAACGACGACAGTGGCACCGCCGACAGAGTCCCTTTTA	1500
Db	1978	AGCAAGGAAATGATTCGGTTCAGAAACGACGACAGTGGCACCGCCGACAGAGTCCCTTTTA	2037
QY	1501	TCCTCATGTCATCGAAGTGGGAGCGCTGAGAGCAATTACACACACAGCTCCCTTTTCGGC	1560
Db	2038	TCCTCATGTCATCGAAGTGGGAGCGCTGAGAGCAATTACACACACAGCTCCCTTTTCGGC	2097
QY	1561	CTTTCCACAGGCGACACAGCACTTCACGAAATCTCTGCTGCGCTTAAAGGCTGGCAC	1620
Db	2098	CTTTCCACAGGCGACACAGCACTTCACGAAATCTCTGCTGCGCTTAAAGGCTGGCAC	2157
QY	1621	TGCGAATCTTGGCCCCCGACAGCCCTAACCCCTTCCCTGACGAGAGCTGTATTTTGGC	1680
Db	2158	TGCGAATCTTGGCCCCCGACAGCCCTAACCCCTTCCCTGACGAGAGCTGTATTTTGGC	2217
QY	1681	ACAGAGTCTCTCACTTCTACTCTGTGCTTACGCCATCTACGAGAGGCAAGTCCAGTTACT	1740
Db	2218	ACAGAGTCTCTCACTTCTACTCTGTGCTTACGCCATCTACGAGAGGCAAGTCCAGTTACT	2277
QY	1741	GCTTACAGCTCAGGCGAGCTGGCCCATCTTGGGAGACCAAGTCTATTTCTGTGCGAGCGG	1800
Db	2278	GCTTACAGCTCAGGCGAGCTGGCCCATCTTGGGAGACCAAGTCTATTTCTGTGCGAGCGG	2337
QY	1801	CAGAAGCCAAGTGCAGAGCTGACTCGCGCGGAGCTGCGCATGAAAGAGAGCCCTTTGAA	1860
Db	2338	CAGAAGCCAAGTGCAGAGCTGACTCGCGCGGAGCTGCGCATGAAAGAGAGCCCTTTGAA	2397
QY	1861	AAGCAGTTTAAAGCGAAGAGCTGCCAATGGAATTTTGAAGAGAGCATGTCAGAGAAC	1920
Db	2398	AAGCAGTTTAAAGCGAAGAGCTGCCAATGGAATTTTGAAGAGAGCATGTCAGAGAAC	2457
QY	1921	AGGTCACGGGANAAGCTGGGGAAAAGTGGGCACTGCTACTCTTTTGGGGCAGATGGAA	1980
Db	2458	AGGTCACGGGANAAGCTGGGGAAAAGTGGGCACTGCTACTCTTTTGGGGCAGATGGAA	2517
QY	1981	ATCATTTAGAGTCTCC	1995
Db	2518	ATCATTTAGAGTCTCC	2532

```

RESULT 3
AAS15768
ID      AAS15768 standard, cDNA; 3059 BP
XX
AC      AAS15768;

```

[illegible]

QY 121 CACATTTGGAGCCATTAAATCACTGCTCCAAAGCTTAAGAGCGAAGTTGCAACAG 180
| | | | |
Db 247 CACATTTGGAGCCATTAAATCACTGCTCCAAAGCTTAAGAGCGAAGTTGCAACAG 306
| | | | |
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCAGACATTCAGCGAAACATAAGTTGACATTGAT 240
| | | | |
Db 307 GACAAAGTGTAAATTAACAGAGCTCATCAGACATTCAGCGAAACATAAGTTGACATTGAT 366
| | | | |
QY 241 TGCAGTACAGAAAGTTGATTTAGATTAAGATCAAAAGCTCCCAAGATTTGCTCTCTTCA 300
| | | | |
Db 367 TGCAGTACAGAAAGTTGATTTAGATTAAGATCAAAAGCTCCCAAGATTTGCTCTCTTCA 426
| | | | |
QY 301 GACTGTTTCTCACTGACTTCTGAGTAAGTGAAGAGACTTCAACTCTGTTCACTG 360
| | | | |
Db 427 GACTGTTTCTCACTGACTTCTGAGTAAGTGAAGAGACTTCAACTCTGTTCACTG 486
| | | | |
QY 361 CTTCGAGTGGGTTTGTCTGAGTTCTCTGTTGTTCCCTGAGCTCTGTGAAGAAATCC 420
| | | | |
Db 487 CTTCGAGTGGGTTTGTCTGAGTTCTCTGTTGTTCCCTGAGCTCTGTGAAGAAATCC 546
| | | | |
QY 421 ACTCTAGTCCCTTACCTGATTTCTCAGCCTTCTTACCTGTCGAACATTTGGCCCAAC 480
| | | | |
Db 547 ACTCTAGTCCCTTACCTGATTTCTCAGCCTTCTTACCTGTCGAACATTTGGCCCAAC 606
| | | | |
QY 481 CGAATTCCTCCATCTTTATCTTGGCTGCGAGAGATGCTTCACAGAGAGCTGATA 540
| | | | |
Db 607 CGAATTCCTCCATCTTTATCTTGGCTGCGAGAGATGCTTCACAGAGAGCTGATA 666
| | | | |
QY 541 CAGCAGAAATGGAGTTGGTTATGTTTAAATGCGAGCTTATCTGTCGAAGCCCTGACTTT 600
| | | | |
Db 667 CAGCAGAAATGGAGTTGGTTATGTTTAAATGCGAGCTTATCTGTCGAAGCCCTGACTTT 726
| | | | |
QY 601 ATCCCGAGATCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
| | | | |
Db 727 ATCCCGAGATCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
| | | | |
QY 661 CCGTGGTTGGAACAATAGATGATTTGATTTGAAGAAAGCAAAAGCTCCCAATGAGTGGTT 720
| | | | |
Db 787 CCGTGGTTGGAACAATAGATGATTTGATTTGAAGAAAGCAAAAGCTCCCAATGAGTGGTT 846
| | | | |
QY 721 CTAGTGACATGTTTGGTGGATCTCCGCTCCGCAACATCGCTATCGCTATCATCATG 780
| | | | |
Db 847 CTAGTGACATGTTTGGTGGATCTCCGCTCCGCAACATCGCTATCGCTATCATCATG 906
| | | | |
QY 781 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAGAAAGCAAAAGCTTACT 840
| | | | |
Db 907 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAGAAAGCAAAAGCTTACT 966
| | | | |
QY 841 ATATCTCAAACTTCAATTTTCTGAGGCAACTCCTGAGCTATGAGAGAAATTAAGAAC 900
| | | | |
Db 967 ATATCTCAAACTTCAATTTTCTGAGGCAACTCCTGAGCTATGAGAGAAATTAAGAAC 1026
| | | | |
QY 901 CAGACTGAGACATCAGGAGCCAAAGAGCAAACTCAAGCTGTCAGCTGAGAGAACCAAT 960
| | | | |
Db 1027 CAGACTGAGACATCAGGAGCCAAAGAGCAAACTCAAGCTGTCAGCTGAGAGAACCAAT 1086
| | | | |
QY 961 GAACCTGTCCCTGTCTGTCTAGAGAGGTGAACAGAAAGCAAGAGCCCTCTCATCCACC 1020
| | | | |
Db 1087 GAACCTGTCCCTGTCTGTCTAGAGAGGTGAACAGAAAGCAAGAGCCCTCTCATCCACC 1146
| | | | |
QY 1021 TGTGCGGACATCTGTCTAGCTCAGAGGCAAGCAAAAGGCGCTGATCCCGCAGAGGTG 1080
| | | | |
Db 1147 TGTGCGGACATCTGTCTAGCTCAGAGGCAAGCAAAAGGCGCTGATCCCGCAGAGGTG 1206
| | | | |
QY 1081 CCGAGCTGCGCAGCTGAGCCCTGCTGTTAGAGAGACAGCCGCTGTATCAGGCGCTC 1140
| | | | |
Db 1207 CCGAGCTGCGCAGCTGAGCCCTGCTGTTAGAGAGACAGCCGCTGTATCAGGCGCTC 1266
| | | | |
QY 1141 AGTGGCTGCACTGTCTCCGACAGAGCTGAGAGACAGCAATTAAGCTTCAAGCTTCTTC 1200
| | | | |
Db 1267 AGTGGCTGCACTGTCTCCGACAGAGCTGAGAGACAGCAATTAAGCTTCAAGCTTCTTC 1326
| | | | |
QY 1201 TCTCTGATATCAAAATGATTTTCAATTTACAGCCAGCATGGAGATCTTAACTAGCTTC 1260
| | | | |

Db 1327 TCTCTGATATCAAAATGATTTATATTAAGCAGATGGCAGATCTTAACTAGCTTC 1386
| | | | |
QY 1261 TCTCTATCAGAGATGCTTTGAAATCTACAAACCTTCCACTCTCTGATGGGACCAAC 1320
| | | | |
Db 1387 TCTCTATCAGAGATGCTTTGAAATCTACAAACCTTCCACTCTCTGATGGGACCAAC 1446
| | | | |
QY 1321 AAGCTATGCGAGTCTCCCTGTCTGAGGAATCTAGGAGAGAGCTCCCGAAACAGTCT 1380
| | | | |
Db 1447 AAGCTATGCGAGTCTCCCTGTCTGAGGAATCTAGGAGAGAGCTCCCGAAACAGTCT 1506
| | | | |
QY 1381 GATTAAGAGAGAACCCAGATCTCCCAAGAGCTGAGAGCCGCAAGCTTTCAGACAGCAG 1440
| | | | |
Db 1507 GATTAAGAGAGAACCCAGATCTCCCAAGAGCTGAGAGCCGCAAGCTTTCAGACAGCAG 1566
| | | | |
QY 1441 AGCAAGCATTTGATTTGGTCAAGAACAGAGAGTGGCACCCGCCAGAGAGTCCCTTTA 1500
| | | | |
Db 1567 AGCAAGCATTTGATTTGGTCAAGAACAGAGAGTGGCACCCGCCAGAGAGTCCCTTTA 1626
| | | | |
QY 1501 TCTCCATGATGAGAGTGGGAGCGTGGAGGACAAATTAACAACAAGCTTCTTTGGC 1560
| | | | |
Db 1627 TCTCCATGATGAGAGTGGGAGCGTGGAGGACAAATTAACAACAAGCTTCTTTGGC 1686
| | | | |
QY 1561 CTTTCCACAGCCAGAGACACTCAGAAAGTCTGCTGAGCTGGGCTTAAAGGCTGGCAC 1620
| | | | |
Db 1687 CTTTCCACAGCCAGAGACACTCAGAAAGTCTGCTGAGCTGGGCTTAAAGGCTGGCAC 1746
| | | | |
QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCTTCCCTGACAGAGAGCTGATTTTGGC 1680
| | | | |
Db 1747 TCGGATATCTTGGCCCCCAGACCTTACCTTCCCTGACAGAGAGCTGATTTTGGC 1806
| | | | |
QY 1681 ACAGATCTCTCAACATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
| | | | |
Db 1807 ACAGATCTCTCAACATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1866
| | | | |
QY 1741 GCCTACAGCTGACAGCAGCTGCCACTTGGCGAGACCAAGTCTATTTCTGCGCAGCGG 1800
| | | | |
Db 1867 GCCTACAGCTGACAGCAGCTGCCACTTGGCGAGACCAAGTCTATTTCTGCGCAGCGG 1926
| | | | |
QY 1801 CAGAGCCAAATGACAGAGCTGACTGCGCGCGAGCTGCGCTATGAAGAGACCCCTTTGAA 1860
| | | | |
Db 1927 CAGAGCCAAATGACAGAGCTGACTGCGCGCGAGCTGCGCTATGAAGAGACCCCTTTGAA 1986
| | | | |
QY 1861 AAGCAGTTTAAAGCAAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTAGAGAAC 1920
| | | | |
Db 1987 AAGCAGTTTAAAGCAAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTAGAGAAC 2046
| | | | |
QY 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGCAGTCACTTATGCTTTTGGGCAAGTGA 1980
| | | | |
Db 2047 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGCAGTCACTTATGCTTTTGGGCAAGTGA 2106
| | | | |
QY 1981 ATCATTGAGGTCTCC 1995
| | | | |
Db 2107 ATCATTGAGGTCTCC 2121
| | | | |
RESULT 4
ABK47596
ID ABK47596 standard; cDNA; 3496 BP.
XX
XX ABK47596;
AC
DT 02-JUL-2002 (first entry)
XX
XX
DE cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.
XX
KW Human: dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; 88.
XX

OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 562..2559
FT /tag= a
FT /product= "Human dual-specificity phosphatase-3 (DSP-16)
protein"
XX MO200226997-A2.
XX
XX 04-APR-2002.
XX
XX 25-SEP-2001; 2001WO-US030124.
XX
XX 26-SEP-2000; 2000US-0235487P.
XX
XX (CEPT)- CEPT-INC.
XX
XX Luche RM, Wei B;
XX
XX WPI: 2002-315802/35.
XX
XX P-PSDB; AAU9156.
XX
XX New DSP-16 polypeptide, useful for identifying modulators of its
XX activity, which can be used in the treatment of disorders such as
XX Duchenne muscular dystrophy, or cancer.
XX
XX Claim 7; Fig 1; 87pp; English.
XX
XX The present invention relates to a new polypeptide, DSP-16, having a 665
XX amino acid sequence, given in the specification, or a variant having at
XX least 50 % identical residues, which retains the ability to
XX dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
XX invention can be used for identifying agents which modulate DSP-16
XX activity, for modulation of a proliferative response in a cell, survival
XX of a cell, or differentiation of a cell. The cell displays contact
XX inhibition of cell growth or anchorage independent growth and may display
XX altered intercellular adhesion. The agent may modulate apoptosis, or the
XX muscular dystrophy, cancer, graft-versus-host disease, autoimmune
XX diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
XX cell proliferation, and cell cycle abnormalities. The present nucleic
XX acid sequence is that of the human dual-specificity phosphatase-3 (DSP-
XX 16) gene located on chromosome 12p. This sequence encodes the human DSP-
XX 16 protein of the invention
SQ
Sequence 3496 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;
Query Match 99.8%; Score 1991.8; DB 6; Length 3496;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGATGATTTGAACTCAATTTTACTGAGAGGTTGGCTCTGCTGAA 60
DB 562 ATGGCCCATGATGATTTGAACTCAATTTTACTGAGAGGTTGGCTCTGCTGAA 621
QY 61 AGTGAACGGAAGAAAGTGTCTAATGATAGCCGCGCATTTTGGAAATCAATACATCC 120
DB 622 AGTGAACGGAAGAAAGTGTCTAATGATAGCCGCGCATTTTGGAAATCAATACATCC 681
QY 121 CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATGAAGGAAGTTGCAACG 180
DB 682 CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATGAAGGAAGTTGCAACG 741
QY 181 GACAAAGTTTATTAACAGAGCTCATCCAGCATTCAGGAAACATAGAGTTGAT 240
DB 742 GACAAAGTTTATTAACAGAGCTCATCCAGCATTCAGGAAACATAGAGTTGAT 801
QY 241 TGCAGTGAAGAGTTTGAATTTAGATCAAGTCCCAAGATGTTGCCCTCTCTTCA 300
DB 802 TGCAGTGAAGAGTTTGAATTTAGATCAAGTCCCAAGATGTTGCCCTCTCTTCA 861
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAAGAGCTTCACTGTTCACTG 360

DB 862 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAAGAGCTTCACTGTTCACTG 921
QY 361 CTTCGAGAGTGGTTTGTGAGTTCTCTGTTTGTCCCTGAGCCCTGTGAAGAAATCC 420
DB 922 CTTCGAGAGTGGTTTGTGAGTTCTCTGTTTGTCCCTGAGCCCTGTGAAGAAATCC 981
QY 421 ACTCTAGTCCCTACCTGCAATTTCTAGCTTGTCTTACCTGTGCAACATTTGGCCACC 480
DB 982 ACTCTAGTCCCTACCTGCAATTTCTAGCTTGTCTTACCTGTGCAACATTTGGCCACC 1041
QY 481 CGAATTTCTCCCAATCTTATCTTGTGCTGCGAGAGATGCTCTCAACAGAGCTGATG 540
DB 1042 CGAATTTCTCCCAATCTTATCTTGTGCTGCGAGAGATGCTCTCAACAGAGCTGATG 1101
QY 541 CAGCAGAAATGGATTTGGTTAATGTTAAATGCCAGTAACTCTGTCAAAAGCTGACTTT 600
DB 1102 CAGCAGAAATGGATTTGGTTAATGTTAATGCCAGTAACTCTGTCAAAAGCTGACTTT 1161
QY 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 660
DB 1162 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 1221
QY 661 CCGTGTGTGACAAATCAAGTAACTTGAATGAAAGAAAGCTTCAATGATGTGT 720
DB 1222 CCGTGTGTGACAAATCAAGTAACTTGAATGAAAGAAAGCTTCAATGATGTGT 781
QY 721 CTAGTGACCTTTAGCTGGATCTCCGCTCCGACCATGCTATGCTGCTATCATG 840
DB 1282 CTAGTGACCTTTAGCTGGATCTCCGCTCCGACCATGCTATGCTGCTATCATG 900
QY 841 ATATCTCCAACTTCAATTTCTGGGCAACTCTGTGACTATGAGAAAGATTAAAGC 1461
DB 1402 ATATCTCCAACTTCAATTTCTGGGCAACTCTGTGACTATGAGAAAGATTAAAGC 1521
QY 901 CAGACTGAGCATGAGGCGCAAGAGCAAACTCAAGCTGTGACTGAGAGGCAAT 960
DB 1462 CAGACTGAGCATGAGGCGCAAGAGCAAACTCAAGCTGTGACTGAGAGGCAAT 1020
QY 961 GAACCTGCTCCGTGTCTCAAGGTTGACAGAAAGGAGAGCCCTCACTGCTCAACC 1080
DB 1522 GAACCTGCTCCGTGTCTCAAGGTTGACAGAAAGGAGAGCCCTCACTGCTCAACC 1140
QY 1021 TGTCCGACTCTGCTACTCTCAAGGAGGAGGCAAGAGCCCTGATCCGCAAGCTG 1181
DB 1582 TGTCCGACTCTGCTACTCTCAAGGAGGAGGCAAGAGCCCTGATCCGCAAGCTG 1241
QY 1081 CCGAGCGTGCAGCGCTGAGCGCTGCTGTTAGAGACAGCCCGCTGTACAGGCTG 1140
DB 1442 CCGAGCGTGCAGCGCTGAGCGCTGCTGTTAGAGACAGCCCGCTGTACAGGCTG 1200
QY 1141 AGTGGCTGCACCTGTCCGACAGAGCTGGAAGAGCAATAGCTTCAAGCTTCTTC 1260
DB 1702 AGTGGCTGCACCTGTCCGACAGAGCTGGAAGAGCAATAGCTTCAAGCTTCTTC 1320
QY 1201 TCTCTGATATCAATCACTGTTTCAATTTCAAGCGAGATGAGCATCTTACATGCTTC 1260
DB 1762 TCTCTGATATCAATCACTGTTTCAATTTCAAGCGAGATGAGCATCTTACATGCTTC 1320
QY 1261 TCTCTGATATCAATCACTGTTTCAATTTCAAGCGAGATGAGCATCTTACATGCTTC 1381
DB 1822 TCTCTGATATCAATCACTGTTTCAATTTCAAGCGAGATGAGCATCTTACATGCTTC 1440
QY 1321 AAGCTATGCGAGTCTCCCTGTTGAGAACTATGAGAGCAAGTCTCCGAACAGTCT 1380
DB 1882 AAGCTATGCGAGTCTCCCTGTTGAGAACTATGAGAGCAAGTCTCCGAACAGTCT 1440
QY 1381 GATTAAGAGAGAGCAGATCCCAAGAGCTGAGACCGCAGGCTTTGACAGCCAG 1440

Db		1942	GATTAAGGAGAAGCCAGCATCCCAAGAAGCTGGACAGCCGACGGCCTTCAGACAGCCAG	2001
Qy		1441	AGCAAAGCATTTGCATTTCGTGGTCAGAACCGACGACAGTAGTGACCGCCAAGAGTCCCTTTTA	15007
Db		2002	AGCAAGGCATTGCAATTCCGTGCAGAACCGACGACAGTAGTGACCGCCAAGAGTCCCTTTTA	2061
Qy		1501	TCTCCACTGCATTCGAAGTGGGAGGGGTGAAGACAATTACACACAGCTTCCTTTTGGGC	1560
Db		2062	TCTCCACTGCATTCGAAGTGGGAGGGGTGAAGACAATTACACACAGCTTCCTTTTGGGC	2121
Qy		1561	CTTTTCACACAGCCACGACAGCACCTTCAAGAAAGTCTGTGGCTGTGGCCTTAAAGGGCTGGCAC	1620
Db		2122	CTTTTCACACAGCCACGACAGCACCTTCAAGAAAGTCTGTGGCTGTGGCCTTAAAGGGCTGGCAC	2181
Qy		1621	TCGGATATCTTTGGCCCCCCCAGAACCTTACCCCTTCCTTGACCAAGAGCTGGTATTTGGCC	1680
Db		2182	TCGGATATCTTTGGCCCCCCCAGAACCTTACCCCTTCCTTGACCAAGAGCTGGTATTTGGCC	2241
Qy		1681	ACAGAGTCTCTACACTTCTACCTCTGCGGCTCAGGCATCTACGGAAGGAGCCAGTTACTCT	1744
Db		2242	ACAGAGTCTCTACACTTCTACCTCTGCGGCTCAGGCATCTACGGAAGGAGCCAGTTACTCT	2301
Qy		1741	GCCTACAGCTCAGCGCAGCTGCCCACTTGGCGAAGCAAGTCTATTCTGTGGCGAGCGGG	1800
Db		2302	GCCTACAGCTCAGCGCAGCTGCCCACTTGGCGAAGCAAGTCTATTCTGTGGCGAGCGGG	2361
Qy		1801	CAGAAGCCMACTGACAGAGCTGACTCGCGCGCGAGCTGGCATGAAGAGAGCCCTTTGAA	1860
Db		2362	CAGAAGCCMACTGACAGAGCTGACTCGCGCGCGAGCTGGCATGAAGAGAGCCCTTTGAA	2421
Qy		1861	AAGCAGTTTAAACCGAAGAGCTGCCAATTTGGAAATTTGGAGAGGATCATGTCAAGAAAC	1922
Db		2422	AAGCAGTTTAAACCGAAGAGCTGCCAATTTGGAAATTTGGAGAGGATCATGTCAAGAAAC	2481
Qy		1921	AGGTCAACGGGAAGAGCTGGGGGAAAGTGGGCAAGTCAAGCTTTTGGGGCAGCATGGAA	1980
Db		2482	AGGTCAACGGGAAGAGCTGGGGGAAAGTGGGCAAGTCAAGCTTTTGGGGCAGCATGGAA	2541
Qy		1981	ATCATTTGAGGTCTCC 1995	
Db		2542	ATCATTTGAGGTCTCC 2556	
<hr/>				
RESULT 5				
ID	AAAS14639			
AAAS14639 standard; cDNA; 3544 BP.				
XX	AAAS14639;			
AC				
XX				
DT	18-DEC-2001 (first entry)			
DE				
XX				
XX	Human CDNA encoding dual specificity phosphatase 21117.			
KW	Human; BR; dual specificity phosphatase 21117; hepatotropic; cytosratic;			
KW	hematopoietic disorder; autoimmune disorder; diabetes mellitus;			
KW	rheumatoid arthritis; multiple sclerosis; Crohn's disease;			
KW	liver disorder; erythroid associated disorder; haemolytic anaemia;			
KW	cellular proliferative; differentiative disorder; leukaemia;			
KW	acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer.			
OS	Homo sapiens.			
XX				
XX				
PH	Key	Location/Qualifiers		
FT	CDS	589..2586		
FT		/tag= a		
FT		/product= "Dual specificity phosphatase 21117"		
XX				
PV	WO200173059-A2.			
XX				
PD	04-OCT-2001.			
XX				
FP	23-MAR-2001; 2001WO-US009477.			
XX				

PR	24-MAR-2000; 2000US-0191858P.
XX	(MILL-) MILLENNIUM PHARM INC.
PA	Meyers RA;
XX	WPI; 2001-611635/70.
PI	P-PSDB; AAU09016.
DR	
XX	
FT	New human dual specificity polypeptides and nucleic acids for diagnosis
PT	of disease and treatment of e.g. liver disorders.
XX	
PS	Claim 1; Fig 1; 143pp; English.
XX	
CC	The invention relates to two novel human dual specificity phosphatases
CC	designated 21117 and 38692, the nucleic acids encoding them (including
CC	fragments, allelic variants, their complements or nucleic acids that
CC	hybridise to them) and antibodies raised against the proteins. The
CC	antibody is useful for detecting the presence of the polypeptide, and the
CC	nucleic acid fragments are useful for detecting the presence of the
CC	nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
CC	antisense sequences) are useful for modulating the activity or expression
CC	of the polypeptide/nucleic acid, useful for the treatment of e.g.
CC	hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
CC	rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
CC	listed in the specification) liver disorders, erythroid associated
CC	disorders (e.g. haemolytic anaemia) cellular proliferative or
CC	differentiative disorders, leukemias (e.g. acute myeloid leukemia),
CC	carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
CC	38692 are also useful for modulating the proliferation, survival,
CC	migration or differentiation of a 38692 or 21117-expressing cell. The
CC	polypeptide and nucleic acids are useful for identifying modulating
CC	agents. The present sequence encodes the dual specificity phosphatase
CC	21117
XX	
SQ	Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
	99.8%; Score 1991.8; DB 5; Length 3544;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	1 ATGGCCCATGATGATGATGGAATCTCAATTGTATCTGAGAGCTTGCGCTCTGCTGGA 60
DB	589 ATGCCCATGATGATGATGGAATCTCAATTGTATCTGAGAGCTTGCGCTCTGCTGGA 648
QY	61 AGTGAACGGAAAAAGTCGCTCAATTGATGACCGGCATTTGGGAATTAATATCATCC 120
DB	649 AGTGAACGGAAAAAGTCGCTCAATTGATGACCGGCATTTGGGAATTAATATCATCC 708
QY	121 CACATTTTGGAAAGCATTAATATCAATGCTCCCAAGCTTATGAAACGAAAGTTGCCAACG 180
DB	709 CACATTTTGGAAAGCATTAATATCAATGCTCCCAAGCTTATGAAACGAAAGTTGCCAACG 768
QY	181 GACAAAGTGTAAATTAACAGAGCTCATCAGATTCACGGAACATTAAGTTGACATTGAT 240
DB	769 GACAAAGTGTAAATTAACAGAGCTCATCAGATTCACGGAACATTAAGTTGACATTGAT 828
QY	241 TGCAGTCAAGAAAGTTTATGTTATGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB	829 TGCAGTCAAGAAAGTTTATGTTATGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 888
QY	301 GACGTTTTCTCATCTGACTTCTCGGTTAAATCGAGAGAAGAGCTTCAACTCTGTTCACCTG 360
DB	889 GACGTTTTCTCATCTGACTTCTCGGTTAAATCGAGAGAAGAGCTTCAACTCTGTTCACCTG 948
QY	361 CTTGACAGGTGGGTTTGTGAAGTTCTCTCGTTGTTTCCCTGGCCCTCTGTGAAGAAAATCC 420
DB	949 CTTGACAGGTGGGTTTGTGAAGTTCTCTCGTTGTTTCCCTGGCCCTCTGTGAAGAAAATCC 1008
QY	421 ACTTAGTCCCTCATCTGACTTCTCAGCCTTGCTTACTGTGTGCAACATTTGGGCAACC 480
DB	1009 ACTTAGTCCCTCATCTGACTTCTCAGCCTTGCTTACTGTGTGCAACATTTGGGCAACC 1068

QY 481 CGAATCTCCCAATCTTATCTTGGCTCCAGCAGATGCTCAACAAGAGCTGATA 540
 Db 1069 CGAATCTCTCCCAATCTTATCTTGGCTCCAGCAGAGATGCTCAACAAGAGCTGATA 1128
 QY 541 CAGCAAAATGGGATTTGTTATGTTAAATGCGAGTATACCTGTCCAAAGCTGACTTT 600
 Db 1129 CAGCAAAATGGGATTTGTTATGTTAAATGCGAGTATACCTGTCCAAAGCTGACTTT 1188
 QY 601 ATCCCGAGTCTCAATTTCCGCGCTGTGTGATGACAGCTTTTGTGAGAAAATTTTG 660
 Db 1189 ATCCCGAGTCTCAATTTCCGCGCTGTGTGATGACAGCTTTTGTGAGAAAATTTTG 1248
 QY 661 CCGTGTGTTGACAAATAGATGATTTGATGAGAAAGCAAAAGCCCTCCAAATGATGTT 720
 Db 1249 CCGTGTGTTGACAAATAGATGATTTGATGAGAAAGCAAAAGCCCTCCAAATGATGTT 1308
 QY 721 CTATGCACTGTTTAAAGTGGGATCTCCCGCTCCGCAACATGCTATGCTCAATCATG 780
 Db 1309 CTATGCACTGTTTAAAGTGGGATCTCCCGCTCCGCAACATGCTATGCTCAATCATG 1368
 QY 781 AAGAGATGGAATGTCCTTATGATGAGCTTACAGATTTGTGAAAGAAAAGCTACT 840
 Db 1369 AAGAGATGGAATGTCCTTATGATGAGCTTACAGATTTGTGAAAGAAAAGCTACT 1428
 QY 841 ATATCTCCAAACTTTTCTGGGCCCACTCTGATCTATGAGAAAGATTTAAGAAC 900
 Db 1429 ATATCTCCAAACTTTTCTGGGCCCACTCTGATCTATGAGAAAGATTTAAGAAC 1488
 QY 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGACCTGTGAGAAAGCCAAAT 960
 Db 1489 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGACCTGTGAGAAAGCCAAAT 1548
 QY 961 GAACCTGTCCCTGCTCTCAAGAGGTGACAGAAAAGCAAGCCCTCACTCCAGCC 1020
 Db 1549 GAACCTGTCCCTGCTCTCTCAAGAGGTGACAGAAAAGCAAGCCCTCACTCCAGCC 1608
 QY 1021 TGTGCGCACTGTGCTACTTCAAGAGCAGAGCAAAAGCCGTGATCCCGCAGCGTG 1080
 Db 1609 TGTGCGCACTGTGCTACTTCAAGAGCAGAGCAAAAGCCGTGATCCCGCAGCGTG 1668
 QY 1081 CCGAGCGTGCCCAAGCTGACAGCGTGCTGTGAGAGCAAGCCCGTGTACAGCGCTC 1140
 Db 1669 CCGAGCGTGCCCAAGCTGACAGCGTGCTGTGAGAGCAAGCCCGTGTACAGCGCTC 1728
 QY 1141 AGTGGCTGTGACCTGTGCGAGACAGAGCTGAGAAAGAGCAATTAAGCTTCTTC 1200
 Db 1729 AGTGGCTGTGACCTGTGCGAGACAGAGCTGAGAAAGAGCAATTAAGCTTCTTC 1788
 QY 1201 TCTCTGATATCAATCAATCAATTCATATGAGCCAGATGAGCAATCTTACATGCTTC 1260
 Db 1789 TCTCTGATATCAATCAATCAATTCATATGAGCCAGATGAGCAATCTTACATGCTTC 1848
 QY 1261 TCTCTATCAGAAAGATCTTTGAAATCTAACAACTTCACTCTGATGAGGACCAAC 1320
 Db 1849 TCTCTATCAGAAAGATCTTTGAAATCTAACAACTTCACTCTGATGAGGACCAAC 1908
 QY 1321 AAGCTATGCAATTTCTCCCTGTTCAGAGCAATTCAGAGCAATCTCCGAAACAGTCTC 1380
 Db 1909 AAGCTATGCAATTTCTCCCTGTTCAGAGCAATTCAGAGCAATCTCCGAAACAGTCTC 1968
 QY 1381 GATTAAGAGAGAGAGAGAGATCCCAAGAGCTGACAGCCGCTTCAGACAGCCAG 2028
 Db 1969 GATTAAGAGAGAGAGAGATCCCAAGAGCTGACAGCCGCTTCAGACAGCCAG 2088
 QY 1441 AGCAAGCATTTGCAATTCGCTGAGAAACAGAGCAAGCTGACAGCCGCTTCAGACAGCCAG 2028
 Db 2029 AGCAAGCATTTGCAATTCGCTGAGAAACAGAGCAAGCTGACAGCCGCTTCAGACAGCCAG 2088
 QY 1501 TCTCCTGCTGATGCAATGAGAGAGAGAGCAATTTACCAACAGCACTTCTTTTCCGC 1560
 Db 2089 TCTCCTGCTGATGCAATGAGAGAGAGAGCAATTTACCAACAGCACTTCTTTTCCGC 2148
 QY 1561 CTTTTCACAGCAGCAGACCTCAAGAGATGCTGCTGCGCTTGAAGGCGCTGAC 1620

Db 2149 CTTTTCACAGCAGCAGACAGACCTCAAGAGTCTGCTGCGCTTGAAGGCTGAC 2208
 QY 1621 TCGGATATCTTGGCCCCCAGACTCTTACCCCTTCCCTGACACAGACTGATATTTGCC 1680
 Db 2209 TCGGATATCTTGGCCCCCAGACTCTTACCCCTTCCCTGACACAGACTGATATTTGCC 2268
 QY 1681 ACAGAGTCTTCACTTCTACCTGCTGAGCCATCTTACAGAGAGAGCTGCACTCT 1740
 Db 2269 ACAGAGTCTTCACTTCTACCTGCTGAGCCATCTTACAGAGAGAGCTGCACTCT 2328
 QY 1741 GCTTACAGTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 Db 2329 GCTTACAGTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2388
 QY 1801 CAGAAAGCAAGTGAACAGAGCTGACCTGCGCGGAGAGCTGCAAGAGAGCCCTTTGAA 1860
 Db 2389 CAGAAAGCAAGTGAACAGAGCTGACCTGCGCGGAGAGCTGCAAGAGAGCCCTTTGAA 2448
 QY 1861 AAGCAATTTAAACGAGAGCTGCAATGATGATTTGAGAGAGATCATGTCAGAGAAC 1920
 Db 2449 AAGCAATTTAAACGAGAGCTGCAATGATGATTTGAGAGAGATCATGTCAGAGAAC 2508
 QY 1921 AGTCAAGGAGAGAGCTGAGGAGAAAGTGGGAGTCACTGCTTGTGCGGACAGATGAA 1980
 Db 2509 AGTCAAGGAGAGAGCTGAGGAGAAAGTGGGAGTCACTGCTTGTGCGGACAGATGAA 2568
 QY 1981 ATCATTTAGGCTTCC 1995
 Db 2569 ATCATTTAGGCTTCC 2583

RESULT 6
 ABK49402
 ID ABK49402 standard; cDNA; 3544 BP.
 XX
 AC ABK49402;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE cDNA encoding human dual specificity phosphatase 21117 protein.
 XX
 KW Human; dual specificity phosphatase 21117; erythroid-related disorder;
 KW haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
 KW erythrocytosis; liver-related disorder; cancer; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT 589..2586
 FT CDS /tag= a
 FT /product= "Human dual specificity phosphatase 21117"
 FT /note= "Specifically claimed in claim 2"
 PN
 PD US2002034807-A1.
 PD 21-MAR-2002.
 PP 23-MAR-2001; 2001US-00816494.
 PR 24-MAR-2000; 2000US-0191858P.
 PA (MEYER) MEYERS R A.
 XX
 PI Meyers RA;
 XX
 DR WPI; 2002-351088/38.
 DR P-PSDB; AAU79929.
 XX
 PT New nucleic acids, designated 38692 and 21117, encoding dual specificity
 PT phosphatases for treating cell proliferation and differentiation
 PT disorders including hematopoietic and erythroid-related disorders and
 PT cancers.

Db 2449 AAGCAGTTTAAAGCAGAGAGCTGCCAATGCAATTGGAGAGAGCATCATGTGACAGAGAC 2508
Qy 1921 AGGTACGCGGAAGAGCTGGCGGAAAGTGGGCACTGCTAGCTTTTGGCGAGCATGGA 1980
Db 2509 AGGTACGCGGAAGAGCTGGCGGAAAGTGGGCACTGCTAGCTTTTGGCGAGCATGGA 2568
Qy 1981 ATCATTGAGAGCTCTCC 1995
Db 2569 ATCATTGAGAGCTCTCC 2583

RESULT 7
ABK14474
ID ABK14474 standard; cDNA, 3766 BP.
XX
AC ABK14474;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human protein phosphatase 7 (PP7) cDNA sequence.
XX
KW Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;
KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
KW melanoma; myeloma sarcoma; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 538..2535
FT /tag=a
FT /product="Protein_phosphatase_7_(PP7)"
XX
PN W0200210363-A2.
XX
PD 07-FEB-2002.
XX
PF 26-JUL-2001; 2001WO-US023716.
XX
PR 28-JUL-2000; 2000US-0221679P.
PR 03-AUG-2000; 2000US-0223272P.
PR 10-AUG-2000; 2000US-0224309P.
PR 18-AUG-2000; 2000US-0226728P.
PR 30-AUG-2000; 2000US-0229254P.
PR 08-SEP-2000; 2000US-0231366P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE,
PI Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalia ALA, Lu DAM,
PI Tribolety CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB,
PI Wala NK, Kearney L;
XX
DR WPI; 2002-188735/24.
DR P-PSDB; AAU5789.
XX
PT New protein phosphatases, useful for diagnosing, treating or preventing
PT immune system disorders (e.g. Crohn's disease), neurological disorders
PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
XX cancers).
XX
PS Claim 5; Page 114-115; 117pp; English.
XX
XX The present invention relates to a new polypeptide, a naturally occurring
XX amino acid sequence at least 95 % identical to it, a biologically active
XX fragment of it or an immunogenic fragment of it. The polypeptides,
XX polynucleotides, agonists and antagonists are useful for diagnosing,
XX treating or preventing disorders associated with aberrant expression of
XX protein phosphatases (PP), particularly immune system disorders e.g.
XX acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
XX asthma or Crohn's disease, neurological disorders e.g. epilepsy,

CC Huntington's disease, dementia or Parkinson's disease, developmental
CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
CC or sarcoma. The present nucleic acid sequence encodes human protein
CC phosphatase 7 (PP7) which is one of several human protein phosphatases
CC (AAU5789-AAU5792) of the invention
XX
SQ Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;
Query Match 99.8%; Score 1991.8; DB 6; Length 3766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCCCATGAGATGATTTGGAATCAATTTGTTACTGAGAGGTGGTGGCTGCTGGA 60
Db 538 ATGGCCCATGAGATGATTTGGAATCAATTTGTTACTGAGAGGTGGTGGCTGCTGGA 597
Qy 61 AGTGAACGGAAGAAAGCTGCTAATTTGATAGCCGGCCATTGTTGAAATACATATCATCC 120
Db 598 AGTGAACGGAAGAAAGCTGCTAATTTGATAGCCGGCCATTGTTGAAATACATATCATCC 657
Qy 121 CACATTTTGAAGCATTAATATCAATGCTCCAGCTTATGAGAGGAGTTGCAACAG 180
Db 658 CACATTTTGAAGCATTAATATCAATGCTCCAGCTTATGAGAGGAGTTGCAACAG 717
Qy 181 GACAAAGTTTATATACAGAGCTCATCCAGCATTCAGGAAACATTAAGTTGACATTTGAT 240
Db 718 GACAAAGTTTATATACAGAGCTCATCCAGCATTCAGGAAACATTAAGTTGACATTTGAT 777
Qy 241 TGCAGTCAGAAAGTTGATTTGATGATGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
Db 778 TGCAGTCAGAAAGTTGATTTGATGATGATGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 837
Qy 301 GACTGTTTCTCACTGATCTTCTGGGTTAACTGGAGAGAGCTTCAACTGTTCACCTG 360
Db 838 GACTGTTTCTCACTGATCTTCTGGGTTAACTGGAGAGAGCTTCAACTGTTCACCTG 897
Qy 361 CTTCAGAGGAGTTTTCGAGGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
Db 898 CTTCAGAGGAGTTTTCGAGGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 957
Qy 421 ACTTAGTCCCTTACCTGATCTTCTGAGCTGTTTACCTGTTTACCTGTTGAGGCAAC 480
Db 958 ACTTAGTCCCTTACCTGATCTTCTGAGCTGTTTACCTGTTTACCTGTTGAGGCAAC 1017
Qy 481 CGAATCTCCCAATCTTATCTTGCTGCGCAGAGATGCTTCAACAAGAGCTGATA 540
Db 1018 CGAATCTCCCAATCTTATCTTGCTGCGCAGAGATGCTTCAACAAGAGCTGATA 1077
Qy 541 CAGCAGATGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 1078 CAGCAGATGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 1138 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1197
Qy 661 CCGTGTGAGAAATACGATGATTTTATGAGAAAGCAAAAGCTCCCAATGATGATGATG 720
Db 1198 CCGTGTGAGAAATACGATGATTTTATGAGAAAGCAAAAGCTCCCAATGATGATGATG 1257
Qy 721 CTAGTGACATGTTTATGCTGGGATCTCCGCTCCGCAACATGCTTATGCTTATCATG 780
Db 1258 CTAGTGACATGTTTATGCTGGGATCTCCGCTCCGCAACATGCTTATGCTTATCATG 1317
Qy 781 AAGAGAGAGCATGCTTTTATGATGAGCTTACAGATTTGTTGAAAGAAAGAAAGCTTAT 840
Db 1318 AAGAGAGAGCATGCTTTTATGATGAGCTTACAGATTTTGAAGAAAGAAAGCTTAT 1377
Qy 841 ATATCTCAAACTTCAATTTTCTGGGCAATCCCTGAGCTATGAGAGAGATTTAAGAAC 900
Db 1378 ATATCTCAAACTTCAATTTTCTGGGCAATCCCTGAGCTATGAGAGAGATTTAAGAAC 1437

QY 901 CAGACTGAGCATCAGGCGCCAAAGAGCAACTCAGAGCTGTGACCTGAGAAAGCCAAAT 960
 DB 1438 CAGACTGAGCATCAGGCGCCAAAGAGCAACTCAGAGCTGTGACCTGAGAAAGCCAAAT 1497
 QY 961 GAACTGTGCTGTGTCTCAGAGGGTGAACAGAAAAGCAGAGCGCCCTCAGTCCACC 1020
 DB 1498 GAACTGTGCTGTGTCTCAGAGGGTGAACAGAAAAGCAGAGCGCCCTCAGTCCACC 1557
 QY 1021 TGTGCGGACTGTGTCTCAGAGGGTGAACAGAAAAGCAGAGCGCCCTCAGTCCACC 1080
 DB 1558 TGTGCGGACTGTGTCTCAGAGGGTGAACAGAAAAGCAGAGCGCCCTCAGTCCACC 1617
 QY 1081 CCCAGCGTGGCCAGCGTGCAGCGCTGCTGTAGAGGACAGCCCGCTGTATACAGCGCTC 1140
 DB 1618 CCCAGCGTGGCCAGCGTGCAGCGCTGCTGTAGAGGACAGCCCGCTGTATACAGCGCTC 1677
 QY 1141 AGTGGGCTGCACCTGTGCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCCTTC 1200
 DB 1678 AGTGGGCTGCACCTGTGCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCCTTC 1737
 QY 1201 TCTCTGATATCAAAATCAGTTTATATTCAGCCAGCATGCGACATCTTACATGCTTC 1260
 DB 1738 TCTCTGATATCAAAATCAGTTTATATTCAGCCAGCATGCGACATCTTACATGCTTC 1797
 QY 1261 TCTCTATCAGAAATGCTTGAATTAATACTAATAAAGCTTCACTCTGTAGTGGAGCCAAAC 1320
 DB 1798 TCTCTATCAGAAATGCTTGAATTAATACTAATAAAGCTTCACTCTGTAGTGGAGCCAAAC 1857
 QY 1321 AAGCTATGCGAGTTCTCCCTGTTAGAGAACTATCGGAGAGACTCCCGAAACAGTCTT 1380
 DB 1858 AAGCTATGCGAGTTCTCCCTGTTAGAGAACTATCGGAGAGACTCCCGAAACAGTCTT 1917
 QY 1381 GATTAAGAGAGAACCCAGCATCCCAAGAAAGCTGACAGCCCGAGGCTTACAGAGCAG 1440
 DB 1918 GATTAAGAGAGAACCCAGCATCCCAAGAAAGCTGACAGCCCGAGGCTTACAGAGCAG 1977
 QY 1441 AGCAGAGATTTGATTCGATTCAGAAACAGCAGAGAGTGGACAGAGTCCCTTTT 1500
 DB 1978 AGCAGAGATTTGATTCGATTCAGAAACAGCAGAGAGTGGACAGAGTCCCTTTT 2037
 QY 1501 TCTCCACTGATCGAAGTGGAGCGTGGAGAGCAATTAACACACAGCTTCTTTTGGC 1560
 DB 2038 TCTCCACTGATCGAAGTGGAGCGTGGAGAGCAATTAACACACAGCTTCTTTTGGC 2097
 QY 1561 CTTTCCACCCAGCCAGCAGCACTCAGAAATCTGTGCTGCTGGAGCTTAAAGGCTGGCAC 1620
 DB 2098 CTTTCCACCCAGCCAGCAGCACTCAGAAATCTGTGCTGCTGGAGCTTAAAGGCTGGCAC 2157
 QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGC 1680
 DB 2158 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGC 2217
 QY 1681 ACAGAGTCTCTACACTTCTACTGTGCTGCTGACCATCTACGAGAGCAGTGCAGTTACTCT 1740
 DB 2218 ACAGAGTCTCTACACTTCTACTGTGCTGCTGACCATCTACGAGAGCAGTGCAGTTACTCT 2277
 QY 1741 GCTTACAGCTGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
 DB 2278 GCTTACAGCTGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2337
 QY 1801 CAGAAAGCAGAGTGAATGCGGCGAGAGCTGCGATGATGAAGAGAGCCCTTTGAA 1860
 DB 2338 CAGAAAGCAGAGTGAATGCGGCGAGAGCTGCGATGATGAAGAGAGCCCTTTGAA 2397
 QY 1861 AAGCAGTTTAAAGCAGAAAGCTGCAAAATGGAATTTGAGAGAGCATCATGTGAGAGAAC 1920
 DB 2398 AAGCAGTTTAAAGCAGAAAGCTGCAAAATGGAATTTGAGAGAGCATCATGTGAGAGAAC 2457
 QY 1921 AGGTTCACGGAAGAGCTGGGGAAGTGGGAGTCACTGCTTCTTGGGAGAGCATGGA 1980
 DB 2458 AGGTTCACGGAAGAGCTGGGGAAGTGGGAGTCACTGCTTCTTGGGAGAGCATGGA 2517
 QY 1981 ATCATTAGAGTCTCC 1995

DB 2518 ATCATTAGAGTCTCC 2532

RESULT 8
 ABN83966
 ID ABN83966 standard; DNA; 4790 BP.
 XX
 AC ABN83966;
 XX
 XX 06-SEP-2002 (first entry)
 XX
 DE Human gene sequence #13.
 XX
 KM Human; brain; tonsil; hippocampus; foetal brain; diagnosis; gene; de.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 184..2181
 FT /tag= a
 PN WO200252005-A1.
 XX
 PD 04-JUL-2002.
 XX
 PF 20-DEC-2001; 2001MO-JP011217.
 XX
 PR 22-DEC-2000; 2000JP-00389742.
 XX
 PA (KAZU-) KAZUSA DNA RES INST FOUND.
 PA (CELE-) CELESTAR LEXICO-SCI LTD.
 XX
 PI Ohara O, Nagase T, Nakajima D;
 XX
 XX WPI; 2002-500762/53.
 DR P-SDB; ABN97946.
 XX
 PT Genes and their expression products cloned from human cDNA libraries for
 PT treatment and diagnosis of diseases associated with their expression.
 XX
 PS Claim 1(a); Page 111-117; 238pp; Japanese.
 XX
 CC The invention relates to DNA encoding polypeptides directly cloned from
 CC cDNA libraries originating in adult whole brain, human tonsil, human
 CC adult hippocampus and human foetal whole brain. Polypeptides and
 CC polynucleotides of the invention may be used in the investigation of
 CC differential expression of the DNA sequences in normal subjects and
 CC disease patients. They may also be used in the production of antibodies,
 CC oligonucleotide probes and DNA chips for diagnosis and identification of
 CC drugs for treatment of diseases with which the DNA sequences are
 CC associated. The sequences given in records ABN83954-ABN83984 represent
 CC human gene sequences of the invention
 XX
 SQ Sequence 4790 BP; 1293 A; 1064 C; 1152 G; 1281 T; 0 U; 0 Other;

Query Match 99.8%; Score 1991.8; DB 6; Length 4790;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGCCATGAGATGATGGAATCTCAATTTGTTACTGAGAGGTGGTGGCTGTGAGAA 60
 DB 184 ATGGGCCATGAGATGATGGAATCTCAATTTGTTACTGAGAGGTGGTGGCTGTGAGAA 243
 QY 61 AGTGAACGGAAGAAAGTGTGCTAATTTGATGAGCGGCAATTTGTGGAATTAATATCATCC 120
 DB 244 AGTGAACGGAAGAAAGTGTGCTAATTTGATGAGCGGCAATTTGTGGAATTAATATCATCC 303
 QY 121 CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATGAAGAGAGTTGCAACAG 180
 DB 304 CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATGAAGAGAGTTGCAACAG 363
 QY 181 GACAAAGTTTAATTACAGAGCTCATCCAGCATTCAGCAAAATTAAGTTGACATTGAT 240

Db 364 GACAAAGTGTATTAACAGAGCTCATCCAGCAATTCAGGAAACATAAGTTGACATTGAT 423
Qy 241 TGGAGTCAGAAAGGTTGTAGTTATTCAGATCAAAAGCTCCAAAGTGTGTGCTCTCTCTCA 300
Db 424 TGGAGTCAGAAAGGTTGTAGTTATTCAGATCAAAAGCTCCAAAGTGTGTGCTCTCTCTCA 483
Qy 301 GACTGTTTTCTCAGTGTACTCTGTGGTAACTGGAGAAAGCTTCAACTGTGTCACTG 360
Db 484 GACTGTTTTCTCAGTGTACTCTGTGGTAACTGGAGAAAGCTTCAACTGTGTCACTG 543
Qy 361 CTGGAGGTTGGTTTGTGCTGAGTTCTCTGTGTTTTCCCTGGCTCTGTGAGAGAAATCC 420
Db 544 CTGGAGGTTGGTTTGTGCTGAGTTCTCTGTGTTTTCCCTGGCTCTGTGAGAGAAATCC 603
Qy 421 ACTTGAATCCCTACCTGCAATTTCTCAGCTTGTGCTTACCTGTTCCAAATGGGCAACC 480
Db 604 ACTTGAATCCCTACCTGCAATTTCTCAGCTTGTGCTTACCTGTTCCAAATGGGCAACC 663
Qy 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGACGAGATGTCTCAACAGAGAGTGATA 540
Db 664 CGAATTTCTCCCAATCTTTATCTTGGCTGCGACGAGATGTCTCAACAGAGAGTGATA 723
Qy 541 CAGCAGAAATGGATTGGTTATGTGTTAAATGCACTATACCTGTCCAAAGCTGACTTT 600
Db 724 CAGCAGAAATGGATTGGTTATGTGTTAAATGCACTATACCTGTGTCCAAAGCTGACTTT 783
Qy 601 ATCCCGAGTCTCATTTCTCGCTGTGCTGTGAATGAACAGCTTTGTGAGAAATTTTG 660
Db 784 ATCCCGAGTCTCATTTCTCGCTGTGCTGTGAATGAACAGCTTTGTGAGAAATTTTG 843
Qy 661 CCGTGTGTGACAAATCGATAGATTCAATGAGAAAGCAAAAGCTTCAATGATGTGT 720
Db 844 CCGTGTGTGACAAATCGATAGATTCAATGAGAAAGCAAAAGCTTCAATGATGTGT 903
Qy 721 CTATGCACTGTTTACGCTGAGATCTCCGCTCGCCACATCGCTATCGCTACATCATG 780
Db 904 CTATGCACTGTTTACGCTGAGATCTCCGCTCGCCACATCGCTATCGCTACATCATG 963
Qy 781 AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAAAAGACCTTACT 840
Db 964 AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAAAAGACCTTACT 1023
Qy 841 ATATCTCCAAATCTTCAATTTCTGTGGGCAATCTCTGACTATGAGAAAGATTAAAGAC 900
Db 1024 ATATCTCCAAATCTTCAATTTCTGTGGGCAATCTCTGACTATGAGAAAGATTAAAGAC 1083
Qy 901 CAGACTGAGCATGAGGGCCAAAGCAAACTCAAGCTGTGCACTCTGAGAAAGCCAAAT 960
Db 1084 CAGACTGAGCATGAGGGCCAAAGCAAACTCAAGCTGTGCACTCTGAGAAAGCCAAAT 1143
Qy 961 GAACTGTCCCTGCTGTCTCAGAGGTGACAGAAAACGAGACGCCCTCAGTCCACC 1020
Db 1144 GAACTGTCCCTGCTGTCTCAGAGGTGACAGAAAACGAGACGCCCTCAGTCCACC 1203
Qy 1021 TGTGCGCATCTGTCTACCTCAGAGGACAGAGCAAAAGGCCCGTGCATCCCGCAAGCTG 1080
Db 1204 TGTGCGCATCTGTCTACCTCAGAGGACAGAGCAAAAGGCCCGTGCATCCCGCAAGCTG 1263
Qy 1081 CCCAGCGTCCAGAGGTCAGCCGCTGTGTTAGAGCAAGCCCGTGTATCAGGGCTC 1140
Db 1264 CCCAGCGTCCAGAGGTCAGCCGCTGTGTTAGAGCAAGCCCGTGTATCAGGGCTC 1323
Qy 1141 AGTGGCTGCACTGTCTCGACAGACAGGCTGAAAGACAGAAATGATCAAGGCTCTTC 1200
Db 1324 AGTGGCTGCACTGTCTCGACAGACAGGCTGAAAGACAGAAATGATCAAGGCTCTTC 1383
Qy 1201 TCTCTGATATCAATCATGTTTCAATATTCAGCCAGATGGCAGCATCTTACATGGCTTC 1260
Db 1384 TCTCTGATATCAATCATGTTTCAATATTCAGCCAGATGGCAGCATCTTACATGGCTTC 1443
Qy 1261 TCCATCATCAAAAGTCTTTGGAATCTACAAACCTTCACTACTGTGAGAGCAAC 1320
Db 1444 TCCATCATCAAAAGTCTTTGGAATCTACAAACCTTCACTACTGTGAGAGCAAC 1503
Qy 1321 AAGCTATGCAAGTTCTCTCCCTGTGAGAACTATCGAGACAGATCCCGAAACAGATCT 1380
Db 1504 AAGCTATGCAAGTTCTCTCCCTGTGAGAACTATCGAGACAGATCCCGAAACAGATCT 1563
Qy 1381 GATTAAGAGAAAGCAGCATCCCAAGAAAGCTGAGACCCGACGCTTCAAGACAG 1440
Db 1564 GATTAAGAGAAAGCAGCATCCCAAGAAAGCTGAGACCCGACGCTTCAAGACAG 1623
Qy 1441 AGCAAGCAATTCATGCTGTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAG 1500
Db 1624 AGCAAGCAATTCATGCTGTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAG 1683
Qy 1501 TCTCCATGTCATGCAATGAGTGGAGGCTGAGAGCAATTAACAACAGCTTCTTTTGGC 1560
Db 1684 TCTCCATGTCATGCAATGAGTGGAGGCTGAGAGCAATTAACAACAGCTTCTTTTGGC 1743
Qy 1561 CTTCACACGAGCCAGAGCACTTCAAGAAATGTGCTGTGGCTTTAAGGGCTGGAC 1620
Db 1744 CTTCACACGAGCCAGAGCACTTCAAGAAATGTGCTGTGGCTTTAAGGGCTGGAC 1803
Qy 1621 TCGAATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTATTTTGC 1680
Db 1804 TCGAATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTATTTTGC 1863
Qy 1681 ACAAGTCTTCAACATTTCTACTGCTGTGAGCATTTACGAGGACAGTGCAGTACTCT 1740
Db 1864 ACAAGTCTTCAACATTTCTACTGCTGTGAGCATTTACGAGGACAGTGCAGTACTCT 1923
Qy 1741 GCTTACAGCTGACAGCCAGCTGCCACTTGTGCGAGACCAAGTATTTGTGCGACGCG 1800
Db 1924 GCTTACAGCTGACAGCCAGCTGCCACTTGTGCGAGACCAAGTATTTGTGCGACGCG 1983
Qy 1801 CAGAAAGCAAGTACAGAGCTGCACTCCGCGGAGCTGTGAGTAAAGAGCCCTTTGAA 1860
Db 1984 CAGAAAGCAAGTACAGAGCTGCACTCCGCGGAGCTGTGAGTAAAGAGAGCCCTTTGAA 2043
Qy 1861 AAGCAGTTTAAACGAGAAAGCTGCAATGAAATTTGAGAGCATGTCAGAGAAC 1920
Db 2044 AAGCAGTTTAAACGAGAAAGCTGCAATGAAATTTGAGAGCATGTCAGAGAAC 2103
Qy 1921 AGTTCACGGGAAAGCTGTGGGAAAGTGGGCACTGCTTTTGTGCGAGCATGAGAA 1980
Db 2104 AGTTCACGGGAAAGCTGTGGGAAAGTGGGCACTGCTTTTGTGCGAGCATGAGAA 2163
Qy 1981 ATCAATTGAGGTCTCC 1995
Db 2164 ATCAATTGAGGTCTCC 2178

Db 1444 TCCATCATCAAAAGTCTTTGGAATCTACAAACCTTCACTACTGTGAGAGCAAC 1503
Qy 1321 AAGCTATGCAAGTTCTCTCCCTGTGAGAACTATCGAGACAGATCCCGAAACAGATCT 1380
Db 1504 AAGCTATGCAAGTTCTCTCCCTGTGAGAACTATCGAGACAGATCCCGAAACAGATCT 1563
Qy 1381 GATTAAGAGAAAGCAGCATCCCAAGAAAGCTGAGACCCGACGCTTCAAGACAG 1440
Db 1564 GATTAAGAGAAAGCAGCATCCCAAGAAAGCTGAGACCCGACGCTTCAAGACAG 1623
Qy 1441 AGCAAGCAATTCATGCTGTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAG 1500
Db 1624 AGCAAGCAATTCATGCTGTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAG 1683
Qy 1501 TCTCCATGTCATGCAATGAGTGGAGGCTGAGAGCAATTAACAACAGCTTCTTTTGGC 1560
Db 1684 TCTCCATGTCATGCAATGAGTGGAGGCTGAGAGCAATTAACAACAGCTTCTTTTGGC 1743
Qy 1561 CTTCACACGAGCCAGAGCACTTCAAGAAATGTGCTGTGGCTTTAAGGGCTGGAC 1620
Db 1744 CTTCACACGAGCCAGAGCACTTCAAGAAATGTGCTGTGGCTTTAAGGGCTGGAC 1803
Qy 1621 TCGAATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTATTTTGC 1680
Db 1804 TCGAATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTATTTTGC 1863
Qy 1681 ACAAGTCTTCAACATTTCTACTGCTGTGAGCATTTACGAGGACAGTGCAGTACTCT 1740
Db 1864 ACAAGTCTTCAACATTTCTACTGCTGTGAGCATTTACGAGGACAGTGCAGTACTCT 1923
Qy 1741 GCTTACAGCTGACAGCCAGCTGCCACTTGTGCGAGACCAAGTATTTGTGCGACGCG 1800
Db 1924 GCTTACAGCTGACAGCCAGCTGCCACTTGTGCGAGACCAAGTATTTGTGCGACGCG 1983
Qy 1801 CAGAAAGCAAGTACAGAGCTGCACTCCGCGGAGCTGTGAGTAAAGAGCCCTTTGAA 1860
Db 1984 CAGAAAGCAAGTACAGAGCTGCACTCCGCGGAGCTGTGAGTAAAGAGAGCCCTTTGAA 2043
Qy 1861 AAGCAGTTTAAACGAGAAAGCTGCAATGAAATTTGAGAGCATGTCAGAGAAC 1920
Db 2044 AAGCAGTTTAAACGAGAAAGCTGCAATGAAATTTGAGAGCATGTCAGAGAAC 2103
Qy 1921 AGTTCACGGGAAAGCTGTGGGAAAGTGGGCACTGCTTTTGTGCGAGCATGAGAA 1980
Db 2104 AGTTCACGGGAAAGCTGTGGGAAAGTGGGCACTGCTTTTGTGCGAGCATGAGAA 2163
Qy 1981 ATCAATTGAGGTCTCC 1995
Db 2164 ATCAATTGAGGTCTCC 2178

RESULT 9
ABV20833
ID ABV20833 standard; cDNA; 5145 BP.
AC
XX
AC
XX
ABV20833;
DP
XX
13-SEP-2002 (first entry)
DE
XX
Human prostate expression marker cDNA 20824.
KW
XX
Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW
XX
pharmacogenomic marker; gene; ss.
OS
XX
Homo sapiens.
PN
XX
W0200160860-A2.
PD
XX
23-AUG-2001.
PF
XX
20-FEB-2001; 2001MO-US005171.
PR
XX
17-FEB-2000; 2000US-0183319P.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy

1429 ATATCTCCAAACCTTTCAGTTGAGAAC 900

QY	901	AAGCTGAGGACATCAAGGGCCAAAGAGGCAAACTCAAGCTGTGCACTCTGAGAAAGCAAT	960
Db	1489	CAGACTGAGACATCAAGGGCCAAAGAGCAAACTCAAGCTGTGCACTCTGAGAAAGCAAT	15448
QY	961	GAACTGTCTCCGTGTCTCAAGAGGGTGAAGAAAAGCGAGACGCCCTCACTGTCACCC	1020
Db	1549	GAACTGTCTCCGTGTCTCAAGAGGGTGAAGAAAAGCGAGACGCCCTCACTGTCACCC	1608
QY	1021	TGTCCGCACTGTGCTACCTCAGAGGACACAGAGCAAAAGCCCTGTGATCCCGCAGCGTG	108
Db	1609	TGTCCGCACTGTGCTACCTCAGAGGACACAGAGCAAAAGCCCTGTGATCCCGCAGCGTG	1666
QY	1081	CCCAAGCTGTCCAGAGGTGACAGCCGTGCGTGTGAAGAGACAGCCCGCTGTACAGGCGCTC	11440
Db	1669	CCCAAGCTGTCCAGAGGTGACAGCCGTGCGTGTGAAGAGACAGCCCGCTGTACAGGCGCTC	1728
QY	1141	AGTGGGCTGACCTGTCCGACAGACAGGCTGGAGAGACAGCAATAGCTCAAGCTTCCTTC	1200
Db	1729	AGTGGGCTGACCTGTCCGACAGACAGGCTGGAGAGACAGCAATAGCTCAAGCTTCCTTC	1788
QY	1201	TCTCTGATATCAATTCAGTTTCATATTCAGCGCAGCATGGAGCATCTCTTCAATGGCTTC	126
Db	1789	TCTCTGATATCAATTCAGTTTCATATTCAGCGCAGCATGGAGCATCTCTTCAATGGCTTC	1848
QY	1261	TCCTCATCAGAAAGATGCTTTGGAACTACTACAAACCTTCACACTCTGTGATGGAGCAAC	1320
Db	1849	TCCTCATCAGAAAGATGCTTTGGAACTACTACAAACCTTCACACTCTGTGATGGAGCAAC	1908
QY	1321	AAGCTATGCCAGTTCTCCCCGTGTTACAGAACTATTTGGAGCAGACTCCGAAAACCACTCT	138
Db	1909	AAGCTATGCCAGTTCTCCCCGTGTTACAGAACTATTTGGAGCAGACTCCGAAAACCACTCT	1968
QY	1381	GATTAAGAGAGAACCCAGCATCCCCAAAGAGCTGACAGCCGACAGGCTTTCAGACAGCCAG	1448
Db	1969	GATTAAGAGAGAACCCAGCATCCCCAAAGAGCTGACAGCCGACAGGCTTTCAGACAGCCAG	2028
QY	1441	AGCAAGCGATTGCAATTGGTCAAGAACAGCAGCAGAGTGACCCGCCAGAGGTCCCTTTTA	1500
Db	2029	AGCAAGCGATTGCAATTGGTCAAGAACAGCAGCAGAGTGACCCGCCAGAGGTCCCTTTTA	2088
QY	1501	TCTCCACATCGATCCGAAGTGGAGAGGTGGAGACAATTACACACAGCTTCCTTTTGGG	156
Db	2089	TCTCCACATCGATCCGAAGTGGAGAGGTGGAGACAATTACACACAGCTTCCTTTTGGG	2148
QY	1561	CTTTTCACACAGCCAGCAGACCTTCACGAAGTCTGTGTGCTGTGGGCTTTAAGGGCTGAC	1620
Db	2149	CTTTTCACACAGCCAGCAGACCTTCACGAAGTCTGTGTGCTGTGGGCTTTAAGGGCTGAC	2208
QY	1621	TCCGATATCTTTGGCCCCCAGACCTTACCCCTTCCTGACCAAGCATGTGTATTTTGGC	1680
Db	2209	TCCGATATCTTTGGCCCCCAGACCTTACCCCTTCCTGACCAAGCATGTGTATTTTGGC	2268
QY	1681	ACAGAGTCTTCACACTTCTACCTGTGCTCCAGCAATCTACGAGAGGACAGTGCAGTTACTCT	1740
Db	2269	ACAGAGTCTTCACACTTCTACCTGTGCTCCAGCAATCTACGAGAGGACAGTGCAGTTACTCT	2328
QY	1741	GCCCTACAGCTGACGCAAGCTGCCCATTTGCCGAGACCAAGTCTATTTCTGTGCGACGGCG	1800
Db	2329	GCCCTACAGCTGACGCAAGCTGCCCATTTGCCGAGACCAAGTCTATTTCTGTGCGACGGCG	2388
QY	1801	CAGAAAGCCAAATGACAGACTGACCTCCGGGGGAGAGCTGGCATTGAAGAGAGCCCTTTGAA	1860
Db	2389	CAGAAAGCCAAATGACAGACTGACCTCCGGGGGAGAGCTGGCATTGAAGAGAGCCCTTTGAA	2448
QY	1861	AAGCAGTTTAAACGCAAGAGCTGCCAAATGGAAATTTGAGAGAGAGCATGATTCAGAGAAC	1920
Db	2449	AAGCAGTTTAAACGCAAGAGCTGCCAAATGGAAATTTGAGAGAGAGCATGATTCAGAGAAC	2508
QY	1921	AGGTTCAGGGAGAGAGCTGGGGAAAGTGGGCACTCAGTCTTACCTTTTCCGGGACGATGGAA	1980
Db	2509	AGGTTCAGGGAGAGAGCTGGGGAAAGTGGGCACTCAGTCTTACCTTTTCCGGGACGATGGAA	2568

QY	1981	ATCATTTGAGGTCCTCC	1995
Db	2569	ATCATTTGAGGTCCTCC	2583

RESULT 11

ID ABV26680 standard; cDNA; 5145 BP.

AC ABV26680;

DT 16-SEP-2002 (first entry)

DE . Human prostate expression marker cDNA 26671.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; drug

xx
xx Homo sapiens

AA WO200160860-A2
PN

23-AUG-2001.

20-FEB-2001; 2001WO-US005171.

17-FEB-2000; 2000US-0183319P.

PR 25-MAY-2000; 2000US-0207454P

PR 18-JUL-2000; 2000US-0219007P.

XX

1 XX
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
8

XX WDT : 2001-66370E/7E
DB

XX
PT Novel associated with cancerous state of

PT for detecting presence of prostate cancer. Stage of prostate cancer, use
PI probate cells and correlating with presence of prostate cancer, use

claim 1: Page 5388-5389: 11750bp: English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC specification or its complement. (I) is useful for: (a) assessing whether

CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

XX
E314E BB- 1400 A- 113E C- 12E3 C- 134E F- 0 Y- 3 O-ho-

Query Match 99 at: Score 1991 8. DB 5: Length 5145

Best Local Similarity 99.96; Pred. NO. 0;
Matches 1993: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGCCCATGAGATGATTGGAACCTCAAATTGTTACTGAGAGGTTGGTGGCTCTGCTGGA 60

Db 589 ATGCCCATGAGATGTTGGACTCAATTGTTACTGAGAGGTTGGTGCTCTGCTGGA 648

61 AGTGAACGAAAGTCTGCTAATTGATAGCCGCCATTCTGGAATACAATACATCC 120

Db 649 AGTGAACGAAAGTCTGCTAATTGATAGCCGCCATTGTGGAATACATACATCC 708

QY 121 CACATTTTGGAGCCATTATATCACTGCTCCAAGCTTATGAAGCGAAGTTGCAACAG 180

XX 20-FEB-2001; 2001WO-US005171.
XX
XX
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous site of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 3451; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (c) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (d) selecting a composition for inhibiting prostate cancer in a patient;
XX (e) assessing the prostate cell carcinogenic potential of a compound; (f)
XX determining whether prostate cancer has metastasized in a patient; (g)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
XX
XX Query Match 99.8%; Score 1991.8; DB 5; Length 5145;
XX Best Local Similarity 99.98; Pred. No. 0;
XX Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 ATGGCCATGAGATGATTTGGAACCAATTGTTACTGAGAGTTGGTGGCTGTGCGAA 60
DB 589 ATGGCCATGAGATGATTTGGAACCAATTGTTACTGAGAGTTGGTGGCTGTGCGAA 648
QY 61 AGTGAACGGAAGAGTGTCTAATTTGATAGCCGGCATTGTGGAATCAATACATCC 120
DB 649 AGTGAACGGAAGAGTGTCTAATTTGATAGCCGGCATTGTGGAATCAATACATCC 708
QY 121 CACATTTTGAAGCCATTATATCACTGCTCAAGTTTGAAGGAGAGTTGCAACAG 180
DB 709 CACATTTTGAAGCCATTATATCACTGCTCAAGTTTGAAGGAGAGTTGCAACAG 768
QY 181 GACAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATTGAT 240
DB 769 GACAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATTGAT 828
QY 241 TCGAGTCAGAGGTTGATTTAAGATCAAGCTCCCAAGATGTGGCTCTCTCTTCA 300
DB 829 TCGAGTCAGAGGTTGATTTAAGATCAAGCTCCCAAGATGTGGCTCTCTCTTCA 888
QY 301 GACTGTTTTCTCACTGACTCTTGGGTTAACTGGAAGAGCTTCAACTGTTCACCTG 360
DB 889 GACTGTTTTCTCACTGACTCTTGGGTTAACTGGAAGAGCTTCAACTGTTCACCTG 948
QY 361 CTTCAGAGTGGTGTGCTGAGTCTCTCGTTGTTCCCTGGCTCTGTGAAGGAAATCC 420
DB 949 CTTCAGAGTGGTGTGCTGAGTCTCTCGTTGTTCCCTGGCTCTGTGAAGGAAATCC 1008
QY 421 ACTCTAGTCCCTTACCTGACATTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCCAA 480
DB 1009 ACTCTAGTCCCTTACCTGACATTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCCAA 1068

QY 481 CGAATTCCTCCAACTCTTATCTTGGCTGCAGAGATGTCCTCAACAGAGCTGATA 540
DB 1069 CGAATTCCTCCAACTCTTATCTTGGCTGCAGAGATGTCCTCAACAGAGCTGATA 1128
QY 541 CACAGAGATGGATGATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 600
DB 1129 CACAGAGATGGATGATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 1188
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 1189 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
QY 661 CCGTGGTGGACAATCAGATGATTTCAATGAGAAAGCAAAAGCTCCCAATGATGATG 720
DB 1249 CCGTGGTGGACAATCAGATGATTTCAATGAGAAAGCAAAAGCTCCCAATGATGATG 1308
QY 721 CTATGCACTGTTTAACTGGGATCTCCGCTCCGCAACATGCTATGCTTATCATCATG 780
DB 1309 CTATGCACTGTTTAACTGGGATCTCCGCTCCGCAACATGCTTATGCTTATCATCATG 1368
QY 781 AAGAGATGAGCATGCTTCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGACTACT 840
DB 1369 AAGAGATGAGCATGCTTCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGACTACT 1428
QY 841 ATATCTCCAACTTCAATTTCTGGGCAACTCTGCACTATGAGAAAGATTAAGAAC 900
DB 1429 ATATCTCCAACTTCAATTTCTGGGCAACTCTGCACTATGAGAAAGATTAAGAAC 1488
QY 901 CAGACTGAGCATCAGAGGCAAAAGCAAACTCACTGCTGCACTGAGAGCAAAAT 960
DB 1489 CAGACTGAGCATCAGAGGCAAAAGCAAACTCACTGCTGCACTGAGAGCAAAAT 1548
QY 961 GAACTGCTCTGCTGCTCTCAGAGGTTGACAGAAAGGAGAGCCCTCACTCAACC 1020
DB 1549 GAACTGCTCTGCTGCTCTCAGAGGTTGACAGAAAGGAGAGCCCTCACTCAACC 1608
QY 1021 TGTGCGGACCTGCTACCTCAGAGGAGAGCAAAAGCCGCTGATCCGCAAGCTG 1080
DB 1609 TGTGCGGACCTGCTACCTCAGAGGAGAGCAAAAGCCGCTGATCCGCAAGCTG 1668
QY 1081 CCCAGGCTGCCAGCTGAGCCGCTGCTGTTAGAGAGACGCCGCTGATCAGAGGCTC 1140
DB 1669 CCCAGGCTGCCAGCTGAGCCGCTGCTGTTAGAGAGACGCCGCTGATCAGAGGCTC 1728
QY 1141 AGTGGCTGCACTGCTGCTGCTGCAAGAGCTGGAAGACAGCAATAGCTCAAGCTTCTTC 1200
DB 1729 AGTGGCTGCACTGCTGCTGCTGCAAGAGCTGGAAGACAGCAATAGCTCAAGCTTCTTC 1788
QY 1201 TCTCTGATATCAATCACTGTTTCAATGATGATGATGATGATGATGATGATGATG 1260
DB 1789 TCTCTGATATCAATCACTGTTTCAATGATGATGATGATGATGATGATGATGATG 1848
QY 1261 TCTCTATCAGAAAGATGCTTGGAAATCACTCAAACTTCACTGATGATGAGCAAC 1320
DB 1849 TCTCTATCAGAAAGATGCTTGGAAATCACTCAAACTTCACTGATGATGAGCAAC 1908
QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATGAGAGCACTCCGAAACCACTCT 1380
DB 1909 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATGAGAGCACTCCGAAACCACTCT 1968
QY 1381 GATTAAG 1440
DB 1969 GATTAAG 2028
QY 1441 AGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 2029 AGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2088
QY 1501 TCTTCATCTGATGAG 1560
DB 2089 TCTTCATCTGATGAG 2148
QY 1561 CTTTCACAG 1620

Db 2149 CTTTCACAGCCAGAGACCTTCAAGAGTCTGCTGGCTTAAAGGCTGGCAC 2208
Qy 1621 TCGGATATTTGGCCCTCCAGACCTTACCCCTTCCGACACAGACCTGATTTTGGCC 1680
Db 2209 TCGGATATTTGGCCCTCCAGACCTTACCCCTTCCGACACAGACCTGATTTTGGCC 2268
Qy 1681 ACAGAGTCTTCACTTCTTACTCTGCTTCCCTTCCGACACAGACCTGATTTTGGCC 2268
Db 2269 ACAGAGTCTTCACTTCTTACTCTGCTTCCCTTCCGACACAGACCTGATTTTGGCC 2328
Qy 1741 GCTTACAGCTGACAGCAGCTGCTCCCACTTGGGAGAGACCAAGTATTTCTGGGAGCGG 1800
Db 2329 GCTTACAGCTGACAGCAGCTGCTCCCACTTGGGAGAGACCAAGTATTTCTGGGAGCGG 2388
Qy 1801 CAGAGCCAGTGAAGAGCTGACCTGCGCGGAGAGCTGCGATGAGAGAGCCCTTTGAA 1860
Db 2389 CAGAGCCAGTGAAGAGCTGACCTGCGCGGAGAGCTGCGATGAGAGAGCCCTTTGAA 2448
Qy 1861 AAGGAGTTTAAAGCGAGAGCTGCCAATGGAATTTGAGAGAGATCATGTGAGAGAC 1920
Db 2449 AAGGAGTTTAAAGCGAGAGCTGCCAATGGAATTTGAGAGAGATCATGTGAGAGAC 2508
Qy 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGCACTGCTTACGTTTGGGAGCATGAA 1980
Db 2509 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGCACTGCTTACGTTTGGGAGCATGAA 2568
Qy 1981 ATCATTTAGAGTCTCC 1995
Db 2569 ATCATTTAGAGTCTCC 2583

RESULT 13

ABV21092
ID ABV21092 standard; cDNA; 5145 BP.

XX ABV21092;

DT 13-SBP-2002 (first entry)

DE Human prostate expression marker cDNA 21083.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN W0200160860-A2.

XX 23-AUG-2001.

FE 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

XX MPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
prostate cells and correlating with presence of prostate cancer, useful
for detecting presence of prostate cancer, stage of prostate cancer.
Claim 1; Page 3485; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC / (I) is also useful as a pharmacodynamic or pharmacogenomic marker
SO

Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 99.8%; Score 1991.8; DB 5; Length 5145;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATTTGAACTCAATTTTACTGAGAGGTGGCTCTGCTGAA 60

Db 589 ATGGCCCATGAGATGATTTGAACTCAATTTTACTGAGAGGTGGCTCTGCTGAA 648

Qy 61 AGTGAACGGAAGAGTGTCTAATTTGATGAGCGGCAATTTGGAGATCAATATCCTC 120

Db 649 AGTGAACGGAAGAGTGTCTAATTTGATGAGCGGCAATTTGGAGATCAATATCCTC 708

Qy 121 CACATTTTGGAGAGCTTAAATATCACTGCTCCAAAGTTTATGAGAGAGTTGCAACAG 180

Db 709 CACATTTTGGAGAGCTTAAATATCACTGCTCCAAAGTTTATGAGAGAGTTGCAACAG 768

Qy 181 GACAAAGTGTAAATACAGAGCTCATCCAGATTCAGGAGAAACATAGATTGATGAT 240

Db 769 GACAAAGTGTAAATACAGAGCTCATCCAGATTCAGGAGAAACATAGATTGATGAT 828

Qy 241 TGCAGTCAGAGAGTGTAAATACAGAGCTCATCCAGATTCAGGAGAAACATAGATTGAT 300

Db 829 TGCAGTCAGAGAGTGTAAATACAGAGCTCATCCAGATTCAGGAGAAACATAGATTGAT 888

Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAAGAGCTTCAACTCTGTCACTG 360

Db 889 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAAGAGCTTCAACTCTGTCACTG 948

Qy 361 CTTCGAGGTGGGTTTGGTGAAGTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTG 420

Db 949 CTTCGAGGTGGGTTTGGTGAAGTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTG 1008

Qy 421 ACTTGAATCCCTTCACTGATCTTCTGGGTAACTGGAAGAGCTTCAACTCTGTCACTG 480

Db 1009 ACTTGAATCCCTTCACTGATCTTCTGGGTAACTGGAAGAGCTTCAACTCTGTCACTG 1068

Qy 481 CGAATTTTCCCAATCTTATCTTGGCTGCGAGAGATGCTTCAACAGAGAGCTGATA 540

Db 1069 CGAATTTTCCCAATCTTATCTTGGCTGCGAGAGATGCTTCAACAGAGAGCTGATA 1128

Qy 541 CAGAGAAATGGAATGTTTATGTTAAATGAGCTTAACTGCTCAAAAGCTGAACTTT 600

Db 1129 CAGAGAAATGGAATGTTTATGTTAAATGAGCTTAACTGCTCAAAAGCTGAACTTT 1188

Qy 601 ATCCCGAGTCTCATTTCTGCGTGGCTGAGATGAGAGCTTTTGGAGAAATTTTG 660

Db 1189 ATCCCGAGTCTCATTTCTGCGTGGCTGAGATGAGAGCTTTTGGAGAAATTTTG 1248

Qy 661 CCGTGTGGAGAAATCAATGATTTTCAATGAGAGAGAGAGCTTCAATGAGAGAGTT 720

Db 1249 CCGTGTGGAGAAATCAATGATTTTCAATGAGAGAGAGAGCTTCAATGAGAGAGTT 1308

Qy 721 CTAGTGAATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780

Db 1309 CTAGTGAATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1368

Qy 781 AAG 840

Db 1369 AAG 1428

QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAATCCTCGSACTATGAGAAAGATTAAAGAAC 900
DB 1429 ATATCTCCAACTTCAATTTTCTGGGCCAATCCTCGSACTATGAGAAAGATTAAAGAAC 1488
QY 901 CAGACTGAGCATCAGGGCCAAAGAGAACTCAAGCTGTGCACTGAGAAAGCCAAAT 960
DB 1489 CAGACTGAGCATCAGGGCCAAAGAGAACTCAAGCTGTGCACTGAGAAAGCCAAAT 1548
QY 961 GAACTGTCTCTGTCTCTCAGAGAGGTGACAGAAAAAGCAGACGCCCTTCAGTCCACC 1020
DB 1549 GAACTGTCTCTGTCTCTCAGAGAGGTGACAGAAAAAGCAGACGCCCTTCAGTCCACC 1608
QY 1021 TGTGCGSACTGTGTACTCTCAGAGAGGCGAGCAAAAGCCCGTGTATCCCGCAGGCTG 1080
DB 1609 TGTGCGSACTGTGTACTCTCAGAGAGGCGAGCAAAAGCCCGTGTATCCCGCAGGCTG 1668
QY 1081 CCCAGCGTCCCGCAGCGTGCAGCCGCTGTAGAGAGACGCCGCTGTACAGGCGCTC 1140
DB 1669 CCCAGCGTCCCGCAGCGTGCAGCCGCTGTAGAGAGACGCCGCTGTACAGGCGCTC 1728
QY 1141 AGTGGGCTGACCTGTCTCGAGACAGGCTGAAAGACAGAAATAGCTCAAGCTTCCTTC 1200
DB 1729 AGTGGGCTGACCTGTCTCGAGACAGGCTGAAAGACAGAAATAGCTCAAGCTTCCTTC 1788
QY 1201 TCTCTGGATATCAAAATCAGTTTCAATATTCAGCCAGCATGTGAGCATGTGCTTC 1260
DB 1789 TCTCTGGATATCAAAATCAGTTTCAATATTCAGCCAGCATGTGAGCATGTGCTTC 1848
QY 1261 TCTCTCATCAGAAAGATGCTTGGAAATCTAATAAATCTTCACTCTGTAGTGGAGCCAA 1320
DB 1849 TCTCTCATCAGAAAGATGCTTGGAAATCTAATAAATCTTCACTCTGTAGTGGAGCCAA 1308
QY 1321 AAGCTATGCAAGTTCTCCCTGTTCAGAACTATCGAGAGACACTCCGAAACCAAGTCT 1380
DB 1909 AAGCTATGCAAGTTCTCCCTGTTCAGAACTATCGAGAGACACTCCGAAACCAAGTCT 1968
QY 1381 GATPAGAGAGAAACCGACATCCCAAGACCTGACAGCCCGCAGGCTTTCAGACAGCAG 1440
DB 1969 GATPAGAGAGAAACCGACATCCCAAGACCTGACAGCCCGCAGGCTTTCAGACAGCAG 2028
QY 1441 AGCAAGCGATTGCTGCTCAGAACCCAGACAGTGGGACCGCCGAGAGTCCCTTTTA 1500
DB 2029 AGCAAGCGATTGCTGCTCAGAACCCAGACAGTGGGACCGCCGAGAGTCCCTTTTA 2088
QY 1501 TCTCCACTGATCGAAGTGGAGGCTGAGAGCAATTAACCAACCAAGCTTCCTTTGCGC 1560
DB 2089 TCTCCACTGATCGAAGTGGAGGCTGAGAGCAATTAACCAACCAAGCTTCCTTTGCGC 2148
QY 1561 CTTTCCACCCAGCCAGACGACCTTCAGAAATGTGTGCTGGGCTTAAAGGCTGCGAC 1620
DB 2149 CTTTCCACCCAGCCAGACGACCTTCAGAAATGTGTGCTGGGCTTAAAGGCTGCGAC 2208
QY 1621 TCGGAATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACAGACAGCTGATTTTGGC 1680
DB 2209 TCGGAATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACAGACAGCTGATTTTGGC 2268
QY 1681 ACAGAGTCTCAGACTTCTACTGTGCTCAGACCATCTAAGAGGAGCAGTGCAGTTACT 1740
DB 2269 ACAGAGTCTCAGACTTCTACTGTGCTCAGACCATCTAAGAGGAGCAGTGCAGTTACT 2328
QY 1741 GCTTACAGCTGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
DB 2329 GCTTACAGCTGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388
QY 1801 CAGAAAGCAAGTGCAGAGCTGACTCGGCGGAGCTGAGTGAAGAGAGAGCCCTTTGAA 1860
DB 2389 CAGAAAGCAAGTGCAGAGCTGACTCGGCGGAGCTGAGTGAAGAGAGAGCCCTTTGAA 2448
QY 1861 AAGCAGTTTAAACGACAGAGCTGCAATGAAATTTTGAAGAGAGCATCATGTGAGAGAC 1920
DB 2449 AAGCAGTTTAAACGACAGAGCTGCAATGAAATTTTGAAGAGAGCATCATGTGAGAGAC 2508

QY 1921 AGGTACGGGAGAGCTGGGGAAAGTGGGAGTCACTAGCTTTTGGGACAGATGAA 1980
DB 2509 AGGTACGGGAGAGCTGGGGAAAGTGGGAGTCACTAGCTTTTGGGACAGATGAA 2568
QY 1981 ATCATTGAGGTCTCC 1995
DB 2569 ATCATTGAGGTCTCC 2583

RESULT 14
ABV21312
ID ABV21312 standard; cDNA; 5145 BP.
XX
AC ABV21312;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 21303.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183119P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3539; 11750dp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 99.8%; Score 1991.8; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCGCCATGAGATGATTGGAATCTCAATTTTACTGAGAGGTTGGTGTCTGCTGGA 60
DB 589 ATGCGCCATGAGATGATTGGAATCTCAATTTTACTGAGAGGTTGGTGTCTGCTGGA 648
QY 61 AGTGAACGGAAGAAAGTGTGCTAATTGATTAAGCGGCAATTTGTGAATATCATATC 120

Db 649 AGTGGAGCGAAAAAGTCTCTTAATTGATAGCCGGCATTTGTGAAATACATATCTCC 708
 QY 121 CACATTTTGAAGCATTAATATATCACTGTCTCAAGCTTATGAAGGAAGTTGCAACG 180
 Db 709 CACATTTTGAAGCATTAATATCACTGTCTCAAGCTTATGAAGGAAGTTGCAACG 768
 QY 181 GACAAAGTTAATTAAGAGCTCATCCAGCATTCGCGAAACATTAAGTTGACATTTGAT 240
 Db 769 GACAAAGTTAATTAAGAGCTCATCCAGCATTCGCGAAACATTAAGTTGACATTTGAT 828
 QY 241 TGCAGTGAAGGTTGTAAGTTACATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 300
 Db 829 TGCAGTGAAGGTTGTAAGTTACATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 888
 QY 301 GACTGTTTCTCACTGTAATCTGAGGTAACAGGGAAGAGTTTAACTCTGTTACCTG 360
 Db 889 GACTGTTTCTCACTGTAATCTGAGGTAACAGGGAAGAGTTTAACTCTGTTACCTG 948
 QY 361 CTGCAAGGTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTAAGGAAGAAATCC 420
 Db 949 CTGCAAGGTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTAAGGAAGAAATCC 1008
 QY 421 ACTAGTCCCTACCTGATTTCTCAGGCTTGCTTACCTGTTGCCAATTTGGCCAAACC 480
 Db 1009 ACTAGTCCCTACCTGATTTCTCAGGCTTGCTTACCTGTTGCCAATTTGGCCAAACC 1068
 QY 481 CGAATTTCTCCCAATCTTTATCTTGTGCTGCGAGGAGATGCTCAACAAGAGCTGATA 540
 Db 1069 CGAATTTCTCCCAATCTTTATCTTGTGCTGCGAGGAGATGCTCAACAAGAGCTGATA 1128
 QY 541 CAGCAGAAATGGGAATGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCTGATCTT 600
 Db 1129 CAGCAGAAATGGGAATGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCTGATCTT 1188
 QY 601 ATCCCGAGTCTCAATTTCTGCTGTGCTGCTGTAATGAAGAGCTTTTGTGAAGAAATTTTG 660
 Db 1189 ATCCCGAGTCTCAATTTCTGCTGTGCTGCTGTAATGAAGAGCTTTTGTGAAGAAATTTTG 1248
 QY 661 CCGTGTGTGAACAATCAGTAGATTTCATTGAGAAGCAAAAGCTCTCCAAATGATGTGTT 720
 Db 1249 CCGTGTGTGAACAATCAGTAGATTTCATTGAGAAGCAAAAGCTCTCCAAATGATGTGTT 1308
 QY 721 CTAGTGAATGTTAGTGTGGAATCTCCGCTCCGCCAATGCTATGAGCTCAATCATG 780
 Db 1309 CTAGTGAATGTTAGTGTGGAATCTCCGCTCCGCCAATGCTATGAGCTCAATCATG 1368
 QY 781 AAGAGATGAGCATGTCTTTAGTAGAAGCTTACAGATTTGGAAGAAAAAGACTTACT 840
 Db 1369 AAGAGATGAGCATGTCTTTAGTAGAAGCTTACAGATTTGGAAGAAAAAGACTTACT 1428
 QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAAAGAAATTAAAGAC 900
 Db 1429 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAAAGAAATTAAAGAC 1488
 QY 901 CAGACTGAGCATCAAGGCGCAAAAGCAAACTCAAGCTGTGCACTGTGGAAGCCAAAT 960
 Db 1489 CAGACTGAGCATCAAGGCGCAAAAGCAAACTCAAGCTGTGCACTGTGGAAGCCAAAT 1548
 QY 961 GAACCTGTCCCTGCTGTCTCAGAGGTTGACAGAAAAAGCAAGCCCTCAGTCCAGCC 1020
 Db 1549 GAACCTGTCCCTGCTGTCTCAGAGGTTGACAGAAAAAGCAAGCCCTCAGTCCAGCC 1608
 QY 1021 TGTGCGAATCTGTACTCTCAGAGGCAAGCAAAAGCCCTGATCCCGCAGCGTG 1080
 Db 1609 TGTGCGAATCTGTACTCTCAGAGGCAAGCAAAAGCCCTGATCCCGCAGCGTG 1668
 QY 1081 CCCAGCGTCCCGCAGCGTGTGTTTGAAGAGCAGCCGCTGTGTGATCAGGCGCTC 1140
 Db 1669 CCCAGCGTCCCGCAGCGTGTGTTTGAAGAGCAGCCGCTGTGTGATCAGGCGCTC 1728
 QY 1141 AGTGGCTGACCTGTCCGAGCAGGCTGGAAGACAGAAATAGCTCAAGCTTCTCTTC 1200

Db 1729 AGTGGCTGACACCTGTCCGACAGACGCTGGAAGACAGCAATAGCTCAAGCTTCTTC 1788
 QY 1201 TCTGTGATATCAATCAGTTTCAATATTCAGCCAGCATGTGGCAGATCTTACATGGCTTC 1260
 Db 1789 TCTGTGATATCAATCAGTTTCAATATTCAGCCAGCATGTGGCAGATCTTACATGGCTTC 1848
 QY 1261 TCTGTGATATCAATCAGTTTCAATATTCAGCCAGCATGTGGCAGATCTTACATGGCTTC 1320
 Db 1849 TCTGTGATATCAATCAGTTTCAATATTCAGCCAGCATGTGGCAGATCTTACATGGCTTC 1908
 QY 1321 AAGCTATGCAAGTTCTCCCTGTTTCAAGAACTATCCGAGAGCATCTCCGAAACAGTCT 1380
 Db 1909 AAGCTATGCAAGTTCTCCCTGTTTCAAGAACTATCCGAGAGCATCTCCGAAACAGTCT 1968
 QY 1381 GATTAAGAGAGAGCAGCATCTCCGAGAGCTGAGAGCAGCCAGGCTTCAAGACCCAG 2028
 Db 1969 GATTAAGAGAGAGCAGCATCTCCGAGAGCTGAGAGCAGCCAGGCTTCAAGACCCAG 2088
 QY 1441 AGCAAGGATTTGATTTGGTCAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 Db 2029 AGCAAGGATTTGATTTGGTCAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2068
 QY 1501 TCTTCAATGATGGAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
 Db 2089 TCTTCAATGATGGAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2148
 QY 1561 CTTTCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
 Db 2149 CTTTCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2208
 QY 1621 TCGGATATCTGAGCCGCCAGACCTCTTACCCCTTCCCTGACCAAGAGAGAGAGAGAG 1680
 Db 2209 TCGGATATCTGAGCCGCCAGACCTCTTACCCCTTCCCTGACCAAGAGAGAGAGAGAG 2268
 QY 1681 ACGAGTCTCTCACTTCTTACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1740
 Db 2269 ACGAGTCTCTCACTTCTTACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2328
 QY 1741 GCTTACAGTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2388
 Db 2329 GCTTACAGTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2448
 QY 1801 CAGAAAGCAGATGACAGAGCTGATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
 Db 2389 CAGAAAGCAGATGACAGAGCTGATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2448
 QY 1861 AAGCAGTTTAAAGCAGAAAGCTGCAATGTGAATTTGAGAGAGATCATGTCAAGAGAC 1920
 Db 2449 AAGCAGTTTAAAGCAGAAAGCTGCAATGTGAATTTGAGAGAGATCATGTCAAGAGAC 2508
 QY 1921 AGGTCAAGGAGAGAGCTGGGAAAGTGGCAGTCAAGTCTTTTGGGAGAGATGGA 1980
 Db 2509 AGGTCAAGGAGAGAGCTGGGAAAGTGGCAGTCAAGTCTTTTGGGAGAGATGGA 2568
 QY 1981 ATCATTTGAGGTCTCC 1995
 Db 2569 ATCATTTGAGGTCTCC 2583

RESULT 15
 ABV21316 standard; cDNA; 5145 BP.
 ID ABV21316.
 AC ABV21316.
 XX
 DT 13-SBP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 21307.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.

XX WO200160860-A2.
PN 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JR;
XX WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 3540; 117509p; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV2213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
XX
XX Query Match 99.8%; Score 1991.8; DB 5; Length 5145;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 1991; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 ATGGCCCATGATGATTTGAACTCAATTTGTTACTGAGAGGTTGTGCTCTGCTGAA 60
DB 589 ATGGCCCATGATGATTTGAACTCAATTTGTTACTGAGAGGTTGTGCTCTGCTGAA 648
QY 61 AGTGAACCGGAAAGAGCTGCTTAATTTGATGACCGGCAATTTGTGGAATCAATACATCC 120
DB 649 AGTGAACCGGAAAGAGCTGCTTAATTTGATGACCGGCAATTTGTGGAATCAATACATCC 708
QY 121 CACATTTTGAAGCATTAATATCACTGCTCCAAAGTTATGAAGCAAGTTGCAACAG 180
DB 709 CACATTTTGAAGCATTAATATCACTGCTCCAAAGTTATGAAGCAAGTTGCAACAG 768
QY 181 GACAAAGTTTAATTAACAGCTCATCAGCAATTCAGCGAAACATAAGTTGACATTGAT 240
DB 769 GACAAAGTTTAATTAACAGCTCATCAGCAATTCAGCGAAACATAAGTTGACATTGAT 828
QY 241 TGAAGTCAGAAAGTTGATTTAGATCAAGCAAGTCCCAAGATGTTGCTCTCTCTTCA 300
DB 829 TGAAGTCAGAAAGTTGATTTAGATCAAGTCCCAAGATGTTGCTCTCTCTCTTCA 888
QY 301 GACTGTTTCTCACTGACTTCTGGGTAAGTGAAGAGAGCTTCACTCTGTTCACTG 360
DB 889 GACTGTTTCTCACTGACTTCTGGGTAAGTGAAGAGAGCTTCACTCTGTTCACTG 948
QY 361 CTTCAGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGAAATCC 420
DB 949 CTTCAGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGAAATCC 1008

QY 421 ACTAGTCCCTCACTGATTTCTCAGCCCTTGCTTACTCTGCAACATTTGGCCCAACC 480
DB 1009 ACTAGTCCCTCACTGATTTCTCAGCCCTTGCTTACTCTGCAACATTTGGCCCAACC 1068
QY 481 CGAATTTCTTCCCAATCTTTATCTTGCTGCTGACGAGATGCTTCAACAGAGCTGATA 540
DB 1069 CGAATTTCTTCCCAATCTTTATCTTGCTGCTGACGAGATGCTTCAACAGAGCTGATA 1128
QY 541 CAGCAGAAATGGAGTTGGTTATGTGTTAAATCCAGCTATCTGTTCCAAAGCTTGACTT 600
DB 1129 CAGCAGAAATGGAGTTGGTTATGTGTTAAATCCAGCTATCTGTTCCAAAGCTTGACTT 1188
QY 601 ATCCCGAGTCTATTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 1189 ATCCCGAGTCTATTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
QY 661 CCGTGTGTCAGCAATCAGTGAATTTTCAATTTGAGAAAGCAAAAGCTTCAATGATGTT 720
DB 1249 CCGTGTGTCAGCAATCAGTGAATTTTCAATTTGAGAAAGCAAAAGCTTCAATGATGTT 1308
QY 721 CTATGTCATCTTTAGTGGGATCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 780
DB 1309 CTATGTCATCTTTAGTGGGATCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 1368
QY 781 AAGAGATGAGCATGCTTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
DB 1369 AAGAGATGAGCATGCTTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1428
QY 841 ATATCTCCAACTTCAATTTTCTGAGGCAATCTCTGCACTATGAGAAAGATTAGAAC 900
DB 1429 ATATCTCCAACTTCAATTTTCTGAGGCAATCTCTGCACTATGAGAAAGATTAGAAC 1488
QY 901 CAGATGAGCATCAGGCGCCAAAGAGCAACTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 1489 CAGATGAGCATCAGGCGCCAAAGAGCAACTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1548
QY 961 GAACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 1549 GAACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1608
QY 1021 TGTGCGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1609 TGTGCGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1668
QY 1081 CCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1669 CCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1141 AGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1729 AGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1788
QY 1201 TCTCTGATATCAATCAGTTTATTTATTTAGCCAGCATGCTGCTTATAGCTTTC 1260
DB 1789 TCTCTGATATCAATCAGTTTATTTATTTAGCCAGCATGCTGCTTATAGCTTTC 1848
QY 1261 TCTCTATCAGAGATGCTTTGAAATCTTCAAACTTCTCACTGCTGATGAGCAAC 1320
DB 1849 TCTCTATCAGAGATGCTTTGAAATCTTCAAACTTCTCACTGCTGATGAGCAAC 1380
QY 1321 AAGCTATGCAAGTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
DB 1909 AAGCTATGCAAGTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1968
QY 1381 GATTAAGAGGAAAGCCAGCATTTCCCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
DB 1969 GATTAAGAGGAAAGCCAGCATTTCCCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2028
QY 1441 AGCAAGCATTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 2029 AGCAAGCATTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2088
QY 1501 TCTCACTGATGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560

```
Db      2089  TCTCCAGTGCATTCGAAGTGGAGGTGGAGACAAATTACCAACGAGCTTCCCTTTGCGC 2148
OY      1561  CTTTCCACAGCCGCGAGCACTTCACGAAGTGTGTGAGCTGTGGCCCTTAAGGGCTGGGAC 1620
Db      2149  CTTTCCACAGCCGCGAGCACTTCACGAAGTGTGTGAGCTGTGGCCCTTAAGGGCTGGGAC 2208
OY      1621  TCGGATATCTTGGGCCCCGAGACCTTACCCCTTCCCTGACCAAGAGCTGTATTGTC 1680
Db      2209  TCGGATATCTTGGGCCCCGAGACCTTACCCCTTCCCTGACCAAGAGCTGTATTGTC 2268
OY      1681  ACAGAGTCTTCAACATTTCTACTGTGCTCAGCCATCTACGAGGCACTGCGAGTTACTCT 1740
Db      2269  ACAGAGTCTTCAACATTTCTACTGTGCTCAGCCATCTACGAGGCACTGCGAGTTACTCT 2328
OY      1741  GCCTACAGCTGCAGCCAGCTGCCCACTTGGCGGAGACCAAGTCTATTGTGCGAGGCGG 1800
Db      2329  GCCTACAGCTGCAGCCAGCTGCCCACTTGGCGGAGACCAAGTCTATTGTGCGAGGCGG 2388
OY      1801  CAGAAAGCCAAAGTACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCTTGAA 1860
Db      2389  CAGAAAGCCAAAGTACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCTTGAA 2448
OY      1861  AAGCAGTTTAAAGCGAGAAGCTGCCAATGGAATTTGAGAGAGCATCATGTCAAGAAC 1920
Db      2449  AAGCAGTTTAAAGCGAGAAGCTGCCAATGGAATTTGAGAGAGCATCATGTCAAGAAC 2508
OY      1921  AGGTACCGGGAAGAGCTGGGGAAGTGGGCACTAGCTTATGCTTTGGGCAAGATGAA 1980
Db      2509  AGGTACCGGGAAGAGCTGGGGAAGTGGGCACTAGCTTATGCTTTGGGCAAGATGAA 2568
OY      1981  ATCATTGAGTCTCC 1995
Db      2569  ATCATTGAGTCTCC 2583
```

Search completed: June 21, 2004, 18:35:44
Job time : 1167.45 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 18:08:06 ; Search time 1191.09 Seconds
(without alignments)
7673.060 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccatgagatgctg9.....tggaaatcatgagctcc 1995

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCRUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1991.8	99.8	1998	9	US-09-816-494-3
2	1991.8	99.8	1998	9	US-10-377-072-27
3	1991.8	99.8	3059	17	US-10-357-026-1
4	1991.8	99.8	3496	9	US-09-964-277-1
5	1991.8	99.8	3544	9	US-09-816-494-1
6	1991.8	99.8	3544	16	US-10-377-072-25
7	1991.8	99.8	3625	13	US-10-425-114-26234
8	1991.8	99.8	3766	13	US-10-343-357-17
9	1991.8	99.8	4790	17	US-10-648-593-115
10	1990.2	99.8	2732	13	US-10-168-506-2
11	1988.6	99.7	2102	16	US-10-094-749-673
12	1988.6	99.7	2966	13	US-10-296-115-520
13	1962.2	98.4	2071	13	US-10-072-012-257
14	1906.8	95.6	2200	13	US-10-072-012-255

15	1653.8	82.9	3332	9	US-09-964-277-20	Sequence 20, Appl
16	879.4	44.1	1916	16	US-10-108-260A-2429	Sequence 2429, Ap
17	382	19.1	2453	14	US-10-005-858-1	Sequence 1, Appl
18	357.8	17.9	2476	13	US-10-220-120-28	Sequence 28, Appl
19	294.2	14.7	2039	13	US-10-072-012-265	Sequence 265, App
20	272.4	13.7	877	9	US-09-764-853-158	Sequence 158, App
21	223	11.2	277	17	US-10-637-855-29	Sequence 29, Appl
22	223	11.2	279	16	US-10-305-720-91	Sequence 91, Appl
23	217	10.9	246	17	US-10-257-026-3	Sequence 3, Appl
24	180.6	9.1	424	9	US-09-783-590-8120	Sequence 8120, Ap
25	124.6	6.2	1337	9	US-09-964-899-46	Sequence 46, Ap
26	124.6	6.2	1449	12	US-10-052-482-132	Sequence 192, App
27	124.6	6.2	1635	17	US-10-648-593-27	Sequence 27, Appl
28	124.6	6.2	1830	15	US-10-346-356-1	Sequence 191, Appl
29	124.6	6.2	1909	13	US-10-052-482-191	Sequence 129, App
30	124.6	6.2	1909	12	US-10-058-270A-129	Sequence 189, Appl
31	124.6	6.2	3286	10	US-09-971-392-9	Sequence 9, Appl
32	115	5.8	1452	12	US-10-052-482-189	Sequence 188, App
33	115	5.8	1977	12	US-10-052-482-188	Sequence 6, Appl
34	110.6	5.5	1356	15	US-10-184-832-6	Sequence 4, Appl
35	110.6	5.5	2741	15	US-10-184-832-4	Sequence 3849, Ap
36	106.8	5.4	2303	9	US-09-880-107-3849	Sequence 3, Appl
37	106.8	5.4	2303	10	US-09-922-146-3	Sequence 603, App
38	106.8	5.4	2303	13	US-10-342-887-603	Sequence 603, App
39	106.8	5.4	2303	13	US-10-172-118-603	Sequence 7, Appl
40	106.8	5.4	2338	15	US-10-363-616-7	Sequence 3, Appl
41	106.6	5.3	1143	13	US-10-184-832-3	Sequence 440, App
42	106.6	5.3	1574	13	US-10-296-115-440	Sequence 1, Appl
43	106.6	5.3	2650	12	US-10-184-832-1	Sequence 2041, Ap
44	103.4	5.2	2104	15	US-10-152-319A-2041	Sequence 654, App
45	103.4	5.2	2104	16	US-10-388-934-654	

ALIGNMENTS

RESULT 1
US-09-816-494-3
Sequence 3, Application US/09816494
Patent No. US20020034807A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OF INVENTION: PHOSPHATE MOLECULES AND USES THEREFOR
TITLE OF INVENTION: PHOSPHATE MOLECULES AND USES THEREFOR
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1998
TYPE: DNA
ORGANISM: Homo sapiens
US-09-816-494-3

Query Match 99.8%; Score 1991.8; DB 9; Length 1998;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGGCCCATGAGTGAATTTGTAAGAGAGGTTGGCTGCTGGAA	60
DB	1	ATGGCCCATGAGTGAATTTGTAAGAGAGGTTGGCTGCTGGAA	60
QY	61	AGTGAACGGAAGAGTCTGCTAATTTGATAGCGGCCATTGTGGAATACATATCC	120
DB	61	AGTGAACGGAAGAGTCTGCTAATTTGATAGCGGCCATTGTGGAATACATATCC	120
QY	121	CACATTTGGAAGCCATTATATCAATGCTCCAGCTTATGAGCGAAGTTGCAACG	180
DB	121	CACATTTGGAAGCCATTATATCAATGCTCCAGCTTATGAGCGAAGTTGCAACG	180

QY 181 GACAAAGTGTATTAACAGAGCTCATCAGCATTCAGGAAACATAAGTTGACATTGAT 240
DB 181 GACAAAGTGTATTAACAGAGCTCATCAGCATTCAGGAAACATAAGTTGACATTGAT 240
QY 241 TGCAGTTCAGAAAGTTGATGTTTAAAGATCAAGGCTCCCAAGATGTGCTCTCTCTTCA 300
DB 241 TGCAGTTCAGAAAGTTGATGTTTAAAGATCAAGGCTCCCAAGATGTGCTCTCTCTTCA 300
QY 301 GACTGTTTTCACCTGATCTTCTGAGTTCTCTGTTTCTCTGAGCTCTGAGGAAATCC 420
DB 301 GACTGTTTTCACCTGATCTTCTGAGTTCTCTGTTTCTCTGAGCTCTGAGGAAATCC 420
QY 361 CTTCAGAGTGGGTTTGTCTGAGTTCTCTGTTTCTCTGAGCTCTGAGGAAATCC 420
DB 361 CTTCAGAGTGGGTTTGTCTGAGTTCTCTGTTTCTCTGAGCTCTGAGGAAATCC 420
QY 421 ACTCTAGTCCCTACCTGACATTTCTCAGGCTCTGATACCTGTGCAATTCCTGAGGCAAC 480
DB 421 ACTCTAGTCCCTACCTGACATTTCTCAGGCTCTGATACCTGTGCAATTCCTGAGGCAAC 480
QY 481 GGAATTTCTTCCCATCTTTATCTTGTGCTGACAGAGATGTCTCAACAGAGCTGATA 540
DB 481 GGAATTTCTTCCCATCTTTATCTTGTGCTGACAGAGATGTCTCAACAGAGCTGATA 540
QY 541 CAGCAGATGGGATTTGATGTTTAAAGTCAAGCTTACCTGTCCAAAGCTGATG 600
DB 541 CAGCAGATGGGATTTGATGTTTAAAGTCAAGCTTACCTGTCCAAAGCTGATG 600
QY 601 ATCCCGAGTCTCATTTTCTGAGTTCTGAGTTGATGACAGCTTTTGTGAAATTTTG 660
DB 601 ATCCCGAGTCTCATTTTCTGAGTTCTGAGTTGATGACAGCTTTTGTGAAATTTTG 660
QY 661 CCGTGGTTCAGAAATCAGTATTTTAAAGTCAAGCTTACCTGTCCAAAGCTGATG 720
DB 661 CCGTGGTTCAGAAATCAGTATTTTAAAGTCAAGCTTACCTGTCCAAAGCTGATG 720
QY 721 CTAGTGCATGTTTACCTGAGATCTCCGCTCCGCAACATCTGATGCTCAATGATG 780
DB 721 CTAGTGCATGTTTACCTGAGATCTCCGCTCCGCAACATCTGATGCTCAATGATG 780
QY 781 AAGAGATGACATGCTTTAGATGAAAGCTTAAAGTTTGTGAAAGAAAGAAAGCTTACT 840
DB 781 AAGAGATGACATGCTTTAGATGAAAGCTTAAAGTTTGTGAAAGAAAGAAAGCTTACT 840
QY 841 ATATCTCAAACTTCAATTTTCTGAGCAATCTCTGAGCTATGAGAAAGATTTAAGAC 900
DB 841 ATATCTCAAACTTCAATTTTCTGAGCAATCTCTGAGCTATGAGAAAGATTTAAGAC 900
QY 901 CAGACTGAGCATCAGGCGCAAGAGCAATCTCAAGCTGTGACCTGAGAAAGCTTAAAT 960
DB 901 CAGACTGAGCATCAGGCGCAAGAGCAATCTCAAGCTGTGACCTGAGAAAGCTTAAAT 960
QY 961 GAACCTGTCTCTGCTCTCAGAGGCTGAGCAAGAAAGAGAGGCTCTCAAGCTTCAACC 1020
DB 961 GAACCTGTCTCTGCTCTCAGAGGCTGAGCAAGAAAGAGAGGCTCTCAAGCTTCAACC 1020
QY 1021 TGTGCGCATCTGTACTCAGAGGCAAGAGCAAGAGGCTCTGATCCCGCAGCGTG 1080
DB 1021 TGTGCGCATCTGTACTCAGAGGCAAGAGCAAGAGGCTCTGATCCCGCAGCGTG 1080
QY 1081 CCCAGGCGCCAGGCTGAGCGCTGCTGTTTAAAGAGAGCGCGCTGAGAGAGGCTCTC 1140
DB 1081 CCCAGGCGCCAGGCTGAGCGCTGCTGTTTAAAGAGAGCGCGCTGAGAGAGGCTCTC 1140
QY 1141 AGTGGGCTGACATGCTGCGAGACAGGCTGAAAGACAGCAATAGCTTCAAGGCTTCTTC 1200
DB 1141 AGTGGGCTGACATGCTGCGAGACAGGCTGAAAGACAGCAATAGCTTCAAGGCTTCTTC 1200
QY 1201 TCTCTGATATCAATATGTTTATATTCAGCAGCATGGAGCATCTTCAATAGGCTTC 1260
DB 1201 TCTCTGATATCAATATGTTTATATTCAGCAGCATGGAGCATCTTCAATAGGCTTC 1260
QY 1261 TCTCATCAGAAATGCTTTGAAATCTCAAACTTCACTTCTGATGGAGCAAC 1320

DB 1261 TCTCATCAGAAATGCTTTGAAATCTCAAACTTCACTTCTGATGGAGCAAC 1320
QY 1321 AAGCTATGCAATGCTTCCCTCTTGAAGATCTTATGAGAGCACTTCCGAAACCAAGCTCT 1380
DB 1321 AAGCTATGCAATGCTTCCCTCTTGAAGATCTTATGAGAGCACTTCCGAAACCAAGCTCT 1380
QY 1381 GATTAAGAGAAAGCCAGCATCCCAAGAGCTGAGACCGCGAGGCTTTTCAAGCAG 1440
DB 1381 GATTAAGAGAAAGCCAGCATCCCAAGAGCTGAGACCGCGAGGCTTTTCAAGCAG 1440
QY 1441 AACAAGCATGATGCTTCTGAGAAACAGAGAGTGCACCGCGAGAGTCCCTTTA 1500
DB 1441 AACAAGCATGATGCTTCTGAGAAACAGAGAGTGCACCGCGAGAGTCCCTTTA 1500
QY 1501 TCTCCATGCAATCGAATGGGAGGCTGAGAGCAATTAACAACAGCTTCTTTTGGC 1560
DB 1501 TCTCCATGCAATCGAATGGGAGGCTGAGAGCAATTAACAACAGCTTCTTTTGGC 1560
QY 1561 CTTCACAGCCAGCAGCACTTCAAGAAATGCTGCTGAGGCTTAAAGGCTGAGC 1620
DB 1561 CTTCACAGCCAGCAGCACTTCAAGAAATGCTGCTGAGGCTTAAAGGCTGAGC 1620
QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCTTCCCTGACAGAGCTGATTTTGGC 1680
DB 1621 TCGGATATCTTGGCCCCCAGACCTTACCTTCCCTGACAGAGCTGATTTTGGC 1680
QY 1681 ACAAGTCTTCACTTCTGCTGCTGAGCATCTTACAGGAGCAGTGCAGTTACTCT 1740
DB 1681 ACAAGTCTTCACTTCTGCTGCTGAGCATCTTACAGGAGCAGTGCAGTTACTCT 1740
QY 1741 GCTTCAAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
DB 1741 GCTTCAAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
QY 1801 CAGAAAGCAAGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
DB 1801 CAGAAAGCAAGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
QY 1861 AAGCAGTTTAAACCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
DB 1861 AAGCAGTTTAAACCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
QY 1921 AGGTACAGGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
DB 1921 AGGTACAGGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
QY 1981 ATCATTTAGGCTTCC 1995
DB 1981 ATCATTTAGGCTTCC 1995

RESULT 2
US-10-377-072-27
Sequence 27, Application US/10377072
Publication No. US20040009501A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals Inc.
APPLICANT: Curtis, Rory A.J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Glucksmann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Myoung
APPLICANT: Tsai, Feng-Ying
TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
FILE REFERENCE: MP103-0180WIM
CURRENT APPLICATION NUMBER: US/10/377,072
CURRENT FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/895,860

;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: US 60/215,370
;; PRIOR FILING DATE: 2000-06-29
;; PRIOR APPLICATION NUMBER: US 09/723,806
;; PRIOR FILING DATE: 2000-11-28
;; PRIOR APPLICATION NUMBER: US 60/187,455
;; PRIOR FILING DATE: 2000-03-07
;; PRIOR APPLICATION NUMBER: US 09/843,297
;; PRIOR FILING DATE: 2001-04-25
;; PRIOR APPLICATION NUMBER: US 60/199,801
;; PRIOR FILING DATE: 2000-04-26
;; PRIOR APPLICATION NUMBER: US 09/861,801
;; PRIOR FILING DATE: 2001-05-21
;; PRIOR APPLICATION NUMBER: US 60/205,508
;; PRIOR FILING DATE: 2000-05-19
;; PRIOR APPLICATION NUMBER: US 09/816,494
;; PRIOR FILING DATE: 2001-03-23
;; PRIOR APPLICATION NUMBER: US 09/815,419
;; PRIOR FILING DATE: 2001-03-22
;; Remaining Prior Application data removed - See File Wrapper or PAM.
;; NUMBER OF SEQ ID NOS: 114
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 27
;; LENGTH: 1998
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1) ... (1998)
US-10-377-072-27

Query Match 99.8%; Score 1991.8; DB 16; Length 1998;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCATGAGATGATGGAACCAATGTTACTAGAGAGTTGTGCTGCTGCGAA 60
DB 1 ATGGCCATGAGATGATGGAACCAATGTTACTAGAGAGTTGTGCTGCTGCGAA 60
QY 61 AGTGAACGGAAGAGTGTCTAATTTAGATAGCCGCGCATTTGTGAAATACATCC 120
DB 61 AGTGAACGGAAGAGTGTCTAATTTAGATAGCCGCGCATTTGTGAAATACATCC 120
QY 121 CACATTTGGAAGCCATTATATCACTGCTCCAAAGTTTGAAGCAAGTTTGCAAC 180
DB 121 CACATTTGGAAGCCATTATATCACTGCTCCAAAGTTTGAAGCAAGTTTGCAAC 180
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGATTCAGCGAAACATAAGTTGACAT 240
DB 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGATTCAGCGAAACATAAGTTGACAT 240
QY 241 TGCAGTCAGAGGTTGTAGTTTACATCAAGCTCCCAAGATGTGCTCTCTTCA 300
DB 241 TGCAGTCAGAGGTTGTAGTTTACATCAAGCTCCCAAGATGTGCTCTCTTCA 300
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACCTGGAAGAAAGCTTCACTCTGTTCA 360
DB 301 GACTGTTTCTCACTGATCTTCTGGGTAACCTGGAAGAAAGCTTCACTCTGTTCA 360
QY 361 CTTCGAGAGTGGGTTGTGATGCTCTGCTGTTTCCCTGCGCTCTGTGAAGAAATCC 420
DB 361 CTTCGAGAGTGGGTTGTGATGCTCTGCTGTTTCCCTGCGCTCTGTGAAGAAATCC 420
QY 421 ACTCTAGTCCCTACCTGATCTTCTGAGCTTGTCTTACCTGTGCAACATTTGGCCACC 480
DB 421 ACTCTAGTCCCTACCTGATCTTCTGAGCTTGTCTTACCTGTGCAACATTTGGCCACC 480
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCAAGGAATGTCCTCAACAGAGCTGATA 540
DB 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCAAGGAATGTCCTCAACAGAGCTGATA 540
QY 541 CAGCAGATGGGATTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCTGACTTT 600
DB 541 CAGCAGATGGGATTTGTTATGTTAAATGCGAGCTATACCTGTGTCCAAAGCTGACTTT 600

DB 541 CAGCAGATGGGATTTGTTATGTTAAATGCGAGCAATACCTGTCCAAAGCTGACTTT 600
QY 601 ATCCCGAGTCTCATTTCTCGGCTGCGCTGGAATGACAGCTTTGTGAAATTTTG 660
DB 601 ATCCCGAGTCTCATTTCTCGGCTGCGCTGGAATGACAGCTTTGTGAAATTTTG 660
QY 661 CCGTGTGGAACAAATCAGTAGATTTCAATGGAAGAAAGCAAGCTTCAATGATGTT 720
DB 661 CCGTGTGGAACAAATCAGTAGATTTCAATGGAAGAAAGCAAGCTTCAATGATGTT 720
QY 721 CTATGCACTGTTTATGCTGGGATCTCCGCTCCGCCAATGCTATCCCTATCATATG 780
DB 721 CTATGCACTGTTTATGCTGGGATCTCCGCTCCGCCAATGCTATCCCTATCATATG 780
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 840
DB 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 840
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
DB 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
QY 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTGGAGGCAAT 960
DB 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTGGAGGCAAT 960
QY 961 GAACCTGTCTCTGCTCTCTCAAGGGTGAAGAAAGGAGGCGCCCTAGTCCAGCC 1020
DB 961 GAACCTGTCTCTGCTCTCTCAAGGGTGAAGAAAGGAGGCGCCCTAGTCCAGCC 1020
QY 1021 TGTGCGGACTGCTACTCAGAGGAGAGCAAGAGCCGCTGATCCCGCAGGCTG 1080
DB 1021 TGTGCGGACTGCTACTCAGAGGAGAGCAAGAGCCGCTGATCCCGCAGGCTG 1080
QY 1081 CCCAGGCTGCCAGGCTGAGCCGTGCTGTTAGAGAGCAAGCCGCTGTAACAGGCTTC 1140
DB 1081 CCCAGGCTGCCAGGCTGAGCCGTGCTGTTAGAGAGCAAGCCGCTGTAACAGGCTTC 1140
QY 1141 AGTGGCTGACCTGTGCCAGAGCAGGCTGGAAGAGCAATAGCTCAAGGCTTC 1200
DB 1141 AGTGGCTGACCTGTGCCAGAGCAGGCTGGAAGAGCAATAGCTCAAGGCTTC 1200
QY 1201 TCTCTGATATCAATCAGTTTCAATTCAGCAGATGAGCAGATCTTACATGCTTC 1260
DB 1201 TCTCTGATATCAATCAGTTTCAATTCAGCAGATGAGCAGATCTTACATGCTTC 1260
QY 1261 TCTCTATCAGAGATGCTTTGAAATCTAACAACTTCCACTCTGATGGGACCAAC 1320
DB 1261 TCTCTATCAGAGATGCTTTGAAATCTAACAACTTCCACTCTGATGGGACCAAC 1320
QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATGGAGAGACTCCCGAAACAGATCCT 1380
DB 1321 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATGGAGAGACTCCCGAAACAGATCCT 1380
QY 1381 GATPAGAGAGAGCCAGATCCCAAGAGCTCAGACGCGCAGGCTTCAAGAGCAG 1440
DB 1381 GATPAGAGAGAGCCAGATCCCAAGAGCTCAGACGCGCAGGCTTCAAGAGCAG 1440
QY 1441 AGCAAGCGATTTGATTCGCTCAGAACCAAGAGTGGCAAGCGCCCAAGAGTCCCTTTTA 1500
DB 1441 AGCAAGCGATTTGATTCGCTCAGAACCAAGAGTGGCAAGCGCCCAAGAGTCCCTTTTA 1500
QY 1501 TCTTCATGCTATGGAAGTGGAGGAGGAGGAGAAATTAACAACAGGCTTCTTCGCG 1560
DB 1501 TCTTCATGCTATGGAAGTGGAGGAGGAGGAGAAATTAACAACAGGCTTCTTCGCG 1560
QY 1561 CTTTCAACAGCAGAGCACTCAGAAAGTCTGAGCTGGGCTTAAAGGCTGAGC 1620
DB 1561 CTTTCAACAGCAGAGCACTCAGAAAGTCTGAGCTGGGCTTAAAGGCTGAGC 1620
QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTATTTTGGC 1680
DB 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTATTTTGGC 1680

QY 1441 AGAAGGATTTGATTTGGGTTCAGAACGAGAGAGTGGGACCGGCCAGAGTCCCTTTTA 1500
DB 1567 AGCAAGGATTTGATTTGGGTTCAGAACGAGAGAGTGGGACCGGCCAGAGTCCCTTTTA 1626
QY 1501 TCTTCATGATGCAAGTGGGAGCGTGGAGGACAAATTACCAACAGACTTCTTTTGGG 1560
DB 1627 TCTTCATGATGCAAGTGGGAGCGTGGAGGACAAATTACCAACAGACTTCTTTTGGG 1686
QY 1561 CTTTCCACGACGACGACACTTCAAGAGTCTGCTGGCTGGGCTTTAAGGCTGGCAC 1620
DB 1687 CTTTCCACGACGACGACACTTCAAGAGTCTGCTGGCTGGGCTTTAAGGCTGGCAC 1746
QY 1621 TCGGATATCTTGGCCCCCGGACCTTCAACCTTCCCTGACGACGACCTGGATTTTGGC 1680
DB 1747 TCGGATATCTTGGCCCCCGGACCTTCAACCTTCCCTGACGACGACCTGGATTTTGGC 1806
QY 1681 ACAGAGTCTCACAATTTCTACTCTGCTCAGCCATCTACGAGGAGGAGTGCAGTTACTCT 1740
DB 1807 ACAGAGTCTCACAATTTCTACTCTGCTCAGCCATCTACGAGGAGGAGTGCAGTTACTCT 1866
QY 1741 GCCTACAGCTGACGACGCTGCTCCCACTTGGGAGACCAAGTCTATTTCTGCGCAGGCGG 1800
DB 1867 GCCTACAGCTGACGACGCTGCTCCCACTTGGGAGACCAAGTCTATTTCTGCGCAGGCGG 1926
QY 1801 CAGAGGCAAGTACAGAGCTGACTCGCGGCGGAGCTGAGATGAGAGAGCCCTTTGAA 1860
DB 1927 CAGAGGCAAGTACAGAGCTGACTCGCGGCGGAGCTGAGATGAGAGAGCCCTTTGAA 1986
QY 1861 AAGCAGTTTAAAGCAGAGAGCTGCAAAATGGAATTTGGAGAGAGCATGTCAGAGAAC 1920
DB 1987 AAGCAGTTTAAAGCAGAGAGCTGCAAAATGGAATTTGGAGAGAGCATGTCAGAGAAC 2046
QY 1921 AGGTACGAGGAAAGTGGGAAAGTGGGAGTCAAGTCTTTCGGGACAGATGAA 1980
DB 2047 AGGTACGAGGAAAGTGGGAAAGTGGGAGTCAAGTCTTTCGGGACAGATGAA 2106
QY 1981 ATCATTGAGGCTCC 1995
DB 2107 ATCATTGAGGCTCC 2121

RESULT 4

US-09-964-277-1

; Sequence 1, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucite, Ralf M.
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-1

Query Match 99.8%; Score 1991.8; DB 9; Length 3496;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATTTGAACTCAATTTGTTACTGAGAGTTGGTGTCTGTGGAA 60
DB 562 ATGGCCCATGATGATTTGAACTCAATTTGTTACTGAGAGTTGGTGTCTGTGGAA 621
QY 61 AGTGAACGGAAGAGTGTCTGCTTAATGATAGCCGSCATTTTGGAAATCAATACATCC 120
DB 622 AGTGAACGGAAGAGTGTCTGCTTAATGATAGCCGSCATTTTGGAAATCAATACATCC 681

QY 121 CACATTTGGAAAGCATTAATATCACTGTCCAGCTTATGAAAGGAAGTTGCAACAG 180
DB 682 CACATTTGGAAAGCATTAATATCACTGTCCAGCTTATGAAAGGAAGTTGCAACAG 741
QY 181 GACAAAGTTTAATTAACAGAGCTCATCAGCATTTACGCGAAACATAGATTGACATTGAT 240
DB 742 GACAAAGTTTAATTAACAGAGCTCATCAGCATTTACGCGAAACATAGATTGACATTGAT 801
QY 241 TGCAGTCAGAAAGTTGATTTACATCAAAAGTCCCAAGATTTGCTCTCTTCA 300
DB 802 TGCAGTCAGAAAGTTGATTTACATCAAAAGTCCCAAGATTTGCTCTCTTCA 861
QY 301 GACTGTTTCTCATCTGATCTTCTGGGTAACGAGAAAGCTTCAACTGTTCACCTG 360
DB 862 GACTGTTTCTCATCTGATCTTCTGGGTAACGAGAAAGCTTCAACTGTTCACCTG 921
QY 361 CTTCAGAGTGGGTTTGTGAGTTCTCTGTTGTTTCTGCGCTCTGTGAAAGAAATCC 420
DB 922 CTTCAGAGTGGGTTTGTGAGTTCTCTGTTGTTTCTGCGCTCTGTGAAAGAAATCC 981
QY 421 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGTACCTGTTCGCAACATTTGGCCAAAC 480
DB 982 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGTACCTGTTCGCAACATTTGGCCAAAC 1041
QY 481 CGAAATTTCTCCCAATCTTTATCTTGGCTGACAGAGATGTCTCAACAAAGAGCTGATA 540
DB 1042 CGAAATTTCTCCCAATCTTTATCTTGGCTGACAGAGATGTCTCAACAAAGAGCTGATA 1101
QY 541 CAGCAGAAATGGAAATTTGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCTGACTT 600
DB 1102 CAGCAGAAATGGAAATTTGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCTGACTT 1161
QY 601 ATCCCGAGTCTCATTTCTCGCTGTCTGTGATGACAGCTTTTGTGAAATTTTGG 660
DB 1162 ATCCCGAGTCTCATTTCTCGCTGTCTGTGATGACAGCTTTTGTGAAATTTTGG 1221
QY 661 CCGTGTGTGGAACAATCAGATGATTTCAATGAGAAAGCAAAAGCTCCATATGATGTT 720
DB 1222 CCGTGTGTGGAACAATCAGATGATTTCAATGAGAAAGCAAAAGCTCCATATGATGTT 1281
QY 721 CTAGTCACTGTTTACCTGGGATCTCCGCTCCGACCAATCGCTATTCGCTACATCATG 780
DB 1282 CTAGTCACTGTTTACCTGGGATCTCCGCTCCGACCAATCGCTATTCGCTACATCATG 1341
QY 781 AAGAGATGACATGCTTTAATGATGAAAGCTTACAGATTTGTGAAAGAAAGAAACCTTACT 840
DB 1342 AAGAGATGACATGCTTTAATGATGAAAGCTTACAGATTTGTGAAAGAAAGAAACCTTACT 1401
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
DB 1402 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 1461
QY 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGACCTGTGAGAGGCCAAT 960
DB 1462 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGACCTGTGAGAGGCCAAT 1521
QY 961 GAACTGTCTCTGTCTCTCAAGAGGTTGACAGAAAGGAGAGCGCCCTCACTCAACC 1020
DB 1522 GAACTGTCTCTGTCTCTCAAGAGGTTGACAGAAAGGAGAGCGCCCTCACTCAACC 1581
QY 1021 TGTGCGGACCTCTGCTACCTCAGAGGAGAGGCAAAAGGCGGTGATCCGCGCAGCGTG 1080
DB 1582 TGTGCGGACCTCTGCTACCTCAGAGGAGAGGCAAAAGGCGGTGATCCGCGCAGCGTG 1641
QY 1081 CCCAGGCTGCCAGCTGTGAGCGCTGTCTGTTAAGAGACAGCCCGTGTGTACAGGCGCTC 1140
DB 1642 CCCAGGCTGCCAGCTGTGAGCGCTGTCTGTTAAGAGACAGCCCGTGTGTACAGGCGCTC 1701
QY 1141 AGTGGCTGCACTGTCTCCGCAACAGGCTGTGAAAGACGAAATTAAGTCAAGGCTTCTTC 1200
DB 1702 AGTGGCTGCACTGTCTCCGCAACAGGCTGTGAAAGACGAAATTAAGTCAAGGCTTCTTC 1761
QY 1201 TCTCTGATATCAAAATCAGTTTCATTTATGAGCAGATGGCAGCATCTTACATGCTTC 1260


```
Db 1762 TCTGTGATATCAAAATCACTTTTCATATTCAGCAGCAAGCAGCATCTTAACATGGCTTC
QY 1261 TCCCTCATGAGAAGATCTTTGAGATATCAAAACCTTTCACACTACTCTGATGGAGCAAC
Db 1822 TCTCATCTGAGAAAGATCTTTGAAATACAAACCTTTCACACTACTCTGATGGAGCAAC
QY 1321 AAGCTATGCGAATCTTCCCTGTGTGAGAACTATCCGACAGCACTCCGAAACAGTCTCT
Db 1882 AAGCTATGCGAATCTTCCCTGTGTGAGAACTATCCGACAGCACTCCGAAACAGTCTCT
QY 1381 GATTAGAGGAAGGACGATCTCCCAAGACCTGACAGCCGAGGCTTCAACAGCCAG
Db 1942 GATTAGAGGAAGGACGATCTCCCAAGACCTGACAGCCGAGGCTTCAACAGCCAG
QY 1441 AGCAAGCGATTTGCAATTCGCTCAAGACCGAGAGTGGCAACCGCCAGAGGTCCCTTTTA
Db 2002 AGCAAGCGATTTGCAATTCGCTCAAGACCGAGAGTGGCAACCGCCAGAGGTCCCTTTTA
QY 1501 TCTCAGCTGATGCAAGTGGAGAGCTGTGAGAGCAATTAACAACAGCTTCCTTTGGGC
Db 2062 TCTCAGCTGATGCAAGTGGAGAGCTGTGAGAGCAATTAACAACAGCTTCCTTTGGGC
QY 1561 CTTTCCACAGCAGCAGCAGCACTCAAGAGTGTGCTGGCTGGCTTTAAGAGGCTGGCAC
Db 2122 CTTTCCACAGCAGCAGCAGCACTCAAGAGTGTGCTGGCTGGCTTTAAGAGGCTGGCAC
QY 1621 TCGGATATCTGGGCCCCCGAGAGCTGTACCCCTTCCCTGACCAAGCACTGGTATTTGGCC
Db 2182 TCGGATATCTGGGCCCCCGAGAGCTGTACCCCTTCCCTGACCAAGCACTGGTATTTGGCC
QY 1681 ACAGAGTCTTCACTTCTACTCTGCTCAGCAGCACTTACAGAGGAGGCACTGCACTTCT
Db 2242 ACAGAGTCTTCACTTCTACTCTGCTCAGCAGCACTTACAGAGGAGGCACTGCACTTCT
QY 1741 GCTTACAGTGTGAGCAGCTGCTCCCACTTGGAGAGCACTTATTTCTGTGGAGAGGAGG
Db 2302 GCTTACAGTGTGAGCAGCTGCTCCCACTTGGAGAGCACTTATTTCTGTGGAGAGGAGG
QY 1801 CAGAAAGCAAGTGAAGAGTGAAGTCTGCGGCGAGAGCTGGAGTGAAGAGGAGGAGG
Db 2362 CAGAAAGCAAGTGAAGAGTGAAGTCTGCGGCGAGAGCTGGAGTGAAGAGGAGGAGG
QY 1861 AAGCAGTTTAAAGCCAAAGCTGCTCAATGGAATTTTGAAGAGCATGATGAGAGAAC
Db 2422 AAGCAGTTTAAAGCCAAAGCTGCTCAATGGAATTTTGAAGAGCATGATGAGAGAAC
QY 1921 AAGTCAAGGAGAGAGTGGGGAAGTGGGAGTCACTAGCTTTTGGGAGAGCATGGA
Db 2482 AAGTCAAGGAGAGAGTGGGGAAGTGGGAGTCACTAGCTTTTGGGAGAGCATGGA
QY 1981 ATCATTTGAGGTCTCC 1995
Db 2542 ATCATTTGAGGTCTCC 2556
```

RESULT 5
US-09-816-494-1

Sequence 1, Application US/09816494
Patent No. US20020034807A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
PRIOR APPLICATION NUMBER: 2001-03-23
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 3544

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (589)...(2583)
US-09-816-494-1

Query Match 99.8%; Score 1991.8; DB 9; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGCCCAATGAGATGATTTGAACTTCAATTTGTTACTGAGAGGTTGGGCTGCTGGAGA 60
Db 589 ATGGCCCAATGAGATGATTTGAACTTCAATTTGTTACTGAGAGGTTGGGCTGCTGGAGA 648
QY 61 AGTGGAAAGGAAAGAGTGTCTTAATTTGAATGAGCCGCTATTGTTGGAATATCAATATCC 120
Db 649 AGTGGAAAGGAAAGAGTGTCTTAATTTGAATGAGCCGCTATTGTTGGAATATCAATATCC 708
QY 121 CACATTTTGAAGGCAATTAATATCAATCTGCTCAAGCTTATGAGCGAAGGTTGCAACAG 180
Db 709 CACATTTTGAAGGCAATTAATATCAATCTGCTCAAGCTTATGAGCGAAGGTTGCAACAG 768
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCACTTGAAGGAAACATTAAGTTGACATTGAT 240
Db 769 GACAAAGTGTAAATTAACAGAGCTCATCCAGCACTTGAAGGAAACATTAAGTTGACATTGAT 828
QY 241 TGCAGTCAAGAGGTTGATTTTACATCAAAAGCTCCAAAGTGTGCTCTCTCTTCA 300
Db 829 TGCAGTCAAGAGGTTGATTTTACATCAAAAGCTCCAAAGTGTGCTCTCTCTCTTCA 888
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGGCTTCACTGTTCACTTG 360
Db 889 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGGCTTCACTGTTCACTTG 948
QY 361 CTTCAGAGTGGGTGCTGAGTCTCTGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 949 CTTCAGAGTGGGTGCTGAGTCTCTGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008
QY 421 ACTTGAATCCCTTACCTGATTTCTCAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 1009 ACTTGAATCCCTTACCTGATTTCTCAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
QY 481 CGAATTTCTCCCAATCTTATTTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 1069 CGAATTTCTCCCAATCTTATTTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1128
QY 541 CAGCAAGATGGATGTTGTTATGTTAAATGCAAGCTTATCCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 1129 CAGCAAGATGGATGTTGTTATGTTAAATGCAAGCTTATCCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 601 ATCCCGAGTCTCAATTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 1189 ATCCCGAGTCTCAATTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
QY 661 CCGTGTGGAACAATCAGTATGATTTCAATGAGAAAGGAAAGGCTTCAATGATGTTT 720
Db 1249 CCGTGTGGAACAATCAGTATGATTTCAATGAGAAAGGAAAGGCTTCAATGATGTTT 1308
QY 721 CTAGTCACTGTTTGTGCTGAGTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 1309 CTAGTCACTGTTTGTGCTGAGTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1368
QY 781 AAGAGATGAGCATGCTTTTATGATGAAGTTACAGATTTGTTGAAAGAAAGAAAGAAAGAAAG 840
Db 1369 AAGAGATGAGCATGCTTTTATGATGAAGTTACAGATTTGTTGAAAGAAAGAAAGAAAGAAAG 900
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1428
Db 1429 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1488
QY 901 CAGACTGAGATCAGGCGCAAGAGCAAACTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
```


1489 CAGACTGAGGATCAGGGCCAAAGAGCAATCAAGCTGCTGACCTGAGGAAGCAAT 1548
QY 961 GAACCTGTCCCTGTGTCTCAAGAGGTGACAGAAAAAGAGAGCCCTCACTGACCC 1020
DB 1549 GAACCTGTCCCTGTGTCTCAAGAGGTGACAGAAAAAGAGAGCCCTCACTGACCC 1608
QY 1021 TGTGCGCACTGTGTCTCAAGAGGTGACAGAAAAAGAGAGCCCTCACTGACCC 1080
DB 1609 TGTGCGCACTGTGTCTCAAGAGGTGACAGAAAAAGAGAGCCCTCACTGACCC 1668
QY 1081 CCCAGCTGCGCAGCGTGCAGCGCTGCTGTAGAGAGAGCCCGCTGTAGAGCGCTC 1140
DB 1669 CCCAGCTGCGCAGCGTGCAGCGCTGCTGTAGAGAGAGCCCGCTGTAGAGCGCTC 1728
QY 1141 AGTGGCTGCACTGTGTGCGAGAGAGCTGAGAGAGCAATTAAGCTTCAAGCTTCT 1200
DB 1729 AGTGGCTGCACTGTGTGCGAGAGAGCTGAGAGAGCAATTAAGCTTCAAGCTTCT 1788
QY 1201 TCTGTGATATCAATAGTTTCAATTCAGCCGAGAGAGAGCAATTAAGCTTCT 1260
DB 1789 TCTGTGATATCAATAGTTTCAATTCAGCCGAGAGAGAGCAATTAAGCTTCT 1848
QY 1261 TCTGTATCAGAAAGATGCTTGAATATCAATTAAGCTTCAATTAAGCTTCAAG 1320
DB 1849 TCTGTATCAGAAAGATGCTTGAATATCAATTAAGCTTCAATTAAGCTTCAAG 1908
QY 1321 AAGCTATGCAAGTTCCTCTGTGAGAGAGCTATGAGAGAGCTTCCGAAGCAAGTCT 1380
DB 1909 AAGCTATGCAAGTTCCTCTGTGAGAGAGCTATGAGAGAGCTTCCGAAGCAAGTCT 1968
QY 1381 GATTAAG 1440
DB 1969 GATTAAG 2028
QY 1441 AGCAAGAGATGATGCTTGTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 2029 AGCAAGAGATGATGCTTGTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2088
QY 1501 TCTTCACTGATGAG 1560
DB 2089 TCTTCACTGATGAG 2148
QY 1561 CTTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1620
DB 2149 CTTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2208
QY 1621 TCGGAATATCTTGGGCCCCCAGACCTTACCTTCCCTGACAGCAGCAGCTGATTTGGC 1680
DB 2209 TCGGAATATCTTGGGCCCCCAGACCTTACCTTCCCTGACAGCAGCAGCTGATTTGGC 2268
QY 1681 ACAAGATCTCTCACTTCTAAGCTTGTGCTGAGCAGCAGCAGCAGCAGCAGTACTCT 1740
DB 2269 ACAAGATCTCTCACTTCTAAGCTTGTGCTGAGCAGCAGCAGCAGCAGCAGTACTCT 2328
QY 1741 GCGTACAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1800
DB 2329 GCGTACAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2388
QY 1801 CAGAGAGCAAGTGAAGAGCTGACTCGCGCGAGAGCTGAGAGAGAGAGAGAGAGAG 1860
DB 2389 CAGAGAGCAAGTGAAGAGCTGACTCGCGCGAGAGCTGAGAGAGAGAGAGAGAGAG 2448
QY 1861 AAGCAGTTTAAACGAGAGAGCTGCAATTTGAGAGAGAGAGAGAGAGAGAGAGAG 1920
DB 2449 AAGCAGTTTAAACGAGAGAGCTGCAATTTGAGAGAGAGAGAGAGAGAGAGAGAG 2508
QY 1921 AGGTCAAGGAG 1980
DB 2509 AGGTCAAGGAG 2568
QY 1981 ATCATTTAGAGTCTCC 1995
DB 2569 ATCATTTAGAGTCTCC 2583

RESULT 6
US-10-377-072-25
Sequence 25, Application US/10377072
Publication No. US2004009501A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals Inc.
APPLICANT: Curtis, Rory A.J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Glucksmann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Miyoung
APPLICANT: Tsai, Fong-Ying
TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: M0103-0180NMIM
CURRENT APPLICATION NUMBER: US/10/377, 072
CURRENT FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/895,860
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215,370
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/723,806
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/187,455
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/843,297
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,801
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: US 09/861,801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205,508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/816,494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/815,419
PRIOR FILING DATE: 2001-03-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PASCSEQ for Windows Version 4.0
SEQ ID NO 25
LENGTH: 3544
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (589) ... (2586)
US-10-377-072-25
Query Match 99.8%; Score 1991.8; DB 16; Length 3544;
Best Local Similarity 99.98%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 60
DB 589 ATGGCCCATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 648
QY 61 AGTGAAGCGAAAGAGTGTCTGCTAATTTGATGAGCGCGCATTTTGTGAAATCAATACATCC 120
DB 649 AGTGAAGCGAAAGAGTGTCTGCTAATTTGATGAGCGCGCATTTTGTGAAATCAATACATCC 708
QY 121 CACATTTGGAAGCATTATATATCAATGCTGCTCAAGCTTATGAAGCGAAGTTGCAACAG 180
DB 709 CACATTTGGAAGCATTATATATCAATGCTGCTCAAGCTTATGAAGCGAAGTTGCAACAG 768
QY 181 GACAAAGTTAATTAACAGAGCTATCCAGCAGATTCAGCGAAGATTAAGTTGACATTTGAT 240
DB 769 GACAAAGTTAATTAACAGAGCTATCCAGCAGATTCAGCGAAGATTAAGTTGACATTTGAT 828

QY 241 TCGAGTGAAGAGTTGTTAGTTTACATCAAAAGCTCCAAAGATGTTGCTCTCTTCA 300
Db 829 TCGAGTGAAGAGTTGTTAGTTTACATCAAAAGCTCCAAAGATGTTGCTCTCTTCA 888
QY 301 GACTGTTTTCTCACTGTAATCTCTGGGTAACTGGAAGAGCTTCAACTCTGTTCACTG 360
Db 889 GACTGTTTTCTCACTGTAATCTCTGGGTAACTGGAAGAGCTTCAACTCTGTTCACTG 948
QY 361 CTGCGAGTGGGTTGCTGATGTTCTCTGTTGTTTCCCTGACCTCTGTAAGGAATAATCC 420
Db 949 CTGCGAGTGGGTTGCTGATGTTCTCTGTTGTTTCCCTGACCTCTGTAAGGAATAATCC 1008
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGTCTTACCTGTTCCAACTTGGGCAACC 480
Db 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGTCTTACCTGTTCCAACTTGGGCAACC 1068
QY 481 CGAATCTTCCCAATCTTAACTTGTGCTGCGCAGCGAGATGCTCTCAAGAGAGCTGATA 540
Db 1069 CGAATCTTCCCAATCTTAACTTGTGCTGCGCAGCGAGATGCTCTCAAGAGAGCTGATA 1128
QY 541 CAGCAGATGGGATGGTATGTTAAATGCGAGCTATACCTGTTCCAAAGCTGACTTT 600
Db 1129 CAGCAGATGGGATGGTATGTTAAATGCGAGCTATACCTGTTCCAAAGCTGACTTT 1188
QY 601 ATCCCGAGTCTCATTTCTCGGCTGCTGTAATGAACAGCTATACCTGTTCCAAAGCTG 660
Db 1189 ATCCCGAGTCTCATTTCTCGGCTGCTGTAATGAACAGCTATACCTGTTCCAAAGCTG 1248
QY 661 CCGTGTGGGCAAAATCAGTAGATTTCAATTGAGAAAGCAAAAGCTCAATGATGTT 720
Db 1249 CCGTGTGGGCAAAATCAGTAGATTTCAATTGAGAAAGCAAAAGCTCAATGATGTT 1308
QY 721 CTAGTGCATGTTTGTGCTGGGATCTCCGCTCCGCACTCGCTATCGCTTACATCATG 780
Db 1309 CTAGTGCATGTTTGTGCTGGGATCTCCGCTCCGCACTCGCTTACATCATG 1368
QY 781 AAGAGATGGGATGTTGTTGATGAAGCTTACAGATTTGTGAAAGAAAAGACTTACT 840
Db 1369 AAGAGATGGGATGTTGTTGATGAAGCTTACAGATTTGTGAAAGAAAAGACTTACT 1428
QY 841 ATATCTCAAACTTCAATTTTCTGGGCACTCTGAGCTATGAGAAAGATTAAAGAC 900
Db 1429 ATATCTCAAACTTCAATTTTCTGGGCACTCTGAGCTATGAGAAAGATTAAAGAC 1488
QY 901 CAGACTGAGATCAGGAGGCAAAAGCAAACTCAAGCTGCTGCACTGGAAGCAAAAT 960
Db 1489 CAGACTGAGATCAGGAGGCAAAAGCAAACTCAAGCTGCTGCACTGGAAGCAAAAT 1548
QY 961 GAACCTGTCTCTGCTCTCAAGGGGTGAGAGAAAGCCCTTCAAGTCAACC 1020
Db 1549 GAACCTGTCTCTGCTCTCAAGGGGTGAGAGAAAGCCCTTCAAGTCAACC 1608
QY 1021 TGTGCGCATCTGCTTCAAGGCAAGAGCAAAAGCCCTGATCCCGCAAGCTG 1080
Db 1609 TGTGCGCATCTGCTTCAAGGCAAGAGCAAAAGCCCTGATCCCGCAAGCTG 1668
QY 1081 CCCAGGTTGCCAGGCTGCAAGCCGTCTGTTAGAGACAGCCCTGATCCCGCAAGCTG 1140
Db 1669 CCCAGGTTGCCAGGCTGCAAGCCGTCTGTTAGAGACAGCCCTGATCCCGCAAGCTG 1728
QY 1141 AGTGGGCTGCACTGCTCGGAGACAGGCTGGAAGACAGCAATAGCTCAAGGCTTCT 1200
Db 1729 AGTGGGCTGCACTGCTCGGAGACAGGCTGGAAGACAGCAATAGCTCAAGGCTTCT 1788
QY 1201 TCTCTGATCAATCAATGTTCAATTTCAAGCAGATGGCAATCTTACATGCTTTC 1260
Db 1789 TCTCTGATCAATCAATGTTCAATTTCAAGCAGATGGCAATCTTACATGCTTTC 1848
QY 1261 TCTCTCATCAAGAAAGTCTTGGAACTACTCAAACTTTCATCTGATGGGCAAC 1320
Db 1849 TCTCTCATCAAGAAAGTCTTGGAACTACTCAAACTTTCATCTGATGGGCAAC 1908

QY 1321 AAGTATGCCAGTTTCTCCCTGTTTCAAGAACTATCGAGACACTCCGAAACAGTCTT 1380
Db 1309 AAGTATGCCAGTTTCTCCCTGTTTCAAGAACTATCGAGACACTCCGAAACAGTCTT 1368
QY 1381 GATTAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGAGACCCGAGGCTTTCAGACAGCAG 1440
Db 1369 GATTAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGAGACCCGAGGCTTTCAGACAGCAG 2028
QY 1441 AGCAAGCATTCATTTCTGCTGAGAAACAGACAGAGTGGCAAGCCGCAAGAGTCCCTTTA 1500
Db 2029 AGCAAGCATTCATTTCTGCTGAGAAACAGACAGAGTGGCAAGCCGCAAGAGTCCCTTTA 2088
QY 1501 TCTTCACTGATTCGAAGTGGAGGCTGAGAGACAAATTACCAACAGCTTCTTTTGGC 1560
Db 2089 TCTTCACTGATTCGAAGTGGAGGCTGAGAGACAAATTACCAACAGCTTCTTTTGGC 2148
QY 1561 CTTTCCACGACGACAGCACTTCAAGAACTGCTGCTGAGGCTTAAAGGCTTGGCAG 1620
Db 2149 CTTTCCACGACGACAGCACTTCAAGAACTGCTGCTGAGGCTTAAAGGCTTGGCAG 2208
QY 1621 TCGGATATCTTGGCCCCAGACCTCTACCCCTTCCGAGCAGAGCTGATTTTGGC 1680
Db 2209 TCGGATATCTTGGCCCCAGACCTCTACCCCTTCCGAGCAGAGCTGATTTTGGC 2268
QY 1681 ACAGAGTCTCAACACTTCTGCTGCTGAGCATCTTACGAGGAGTGGCAAGTCTT 1740
Db 2269 ACAGAGTCTCAACACTTCTGCTGCTGAGCATCTTACGAGGAGTGGCAAGTCTT 2328
QY 1741 GCTTACAGCTGCAAGCCAGCTGCCACTTGGGAGACCAAGCTTATTTGTCGAGCGG 1800
Db 2329 GCTTACAGCTGCAAGCCAGCTGCCACTTGGGAGACCAAGCTTATTTGTCGAGCGG 2388
QY 1801 CAGAAACCAAGTGAAGAGCTGACCTCGGCGGAGCTGSCATTAAGAGAGCCCTTTGAA 1860
Db 2389 CAGAAACCAAGTGAAGAGCTGACCTCGGCGGAGCTGSCATTAAGAGAGCCCTTTGAA 2448
QY 1861 AAGCAGTTTAAACGAGAAAGCTGCAATGAAATTTGAGAGAGCATCTGTCAGAGAAC 1920
Db 2449 AAGCAGTTTAAACGAGAAAGCTGCAATGAAATTTGAGAGAGCATCTGTCAGAGAAC 2508
QY 1921 AGGTACCGGAAAGCTGGGGAAGAGTGGCAAGTCAATGCTTTTGGGCAAGTGA 1980
Db 2509 AGGTACCGGAAAGCTGGGGAAGAGTGGGCAAGTCAATGCTTTTGGGCAAGTGA 2568
QY 1981 ATCATTAAGGCTTCC 1995
Db 2569 ATCATTAAGGCTTCC 2583

RESULT 7
US-10-425-114-26234
/ Sequence 26234, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Ltu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ NUMBER OF SEQ ID NOS: 2003-04-28
/ SEQ ID NO 26234
/ LENGTH: 3625
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4119-028-H6_FLI
US-10-425-114-26234

Query Match 99.8%; Score 1991.8; DB 13; Length 3625;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATTTGGAACCAATTGTTACTGAGAGGTTGTCCTGCTGGAA 60
 DB 692 ATGGCCCATGAGATGATTTGGAACCAATTGTTACTGAGAGGTTGTCCTGCTGGAA 751
 QY 61 AGTGAACGAAAAAGTCTGCTAATTGATAGCCGACATTTGTGAATACAATACATCC 120
 DB 752 AGTGAACGAAAAAGTCTGCTAATTGATAGCCGACATTTGTGAATACAATACATCC 811
 QY 121 CACATTTGGAAGCATTAATATCACTGCTCCAGCTTATGAAGCCAGAGTTGCAACAG 180
 DB 812 CACATTTGGAAGCATTAATATCACTGCTCCAGCTTATGAAGCCAGAGTTGCAACAG 871
 QY 181 GACAAAGTGTAAATTAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 240
 DB 872 GACAAAGTGTAAATTAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 931
 QY 241 TGCAGTCAGAAAGTTGATTTAGATCAAAAGCTCCAGAGATGTTGCTCTCTCTTCA 300
 DB 932 TGCAGTCAGAAAGTTGATTTAGATCAAAAGCTCCAGAGATGTTGCTCTCTCTTCA 991
 QY 301 GACTGTTTTCTCACTGACTTCTGCGGTAACTGGAAGAAAGCTTCACTCTTCACTG 360
 DB 992 GACTGTTTTCTCACTGACTTCTGCGGTAACTGGAAGAAAGCTTCACTCTTCACTG 1051
 QY 361 CTTCGAGAGTGGGTTTGTGAGATTCTCGTTGTTTCCCTGGACCTCTGTGAAGAAATCC 420
 DB 1052 CTTCGAGAGTGGGTTTGTGAGATTCTCGTTGTTTCCCTGGACCTCTGTGAAGAAATCC 1111
 QY 421 ACTCTAGTCCCTACCTGACATTTCTCAAGCTTGTCTTACCTGTGCAACATTTGGCCACC 480
 DB 1112 ACTCTAGTCCCTACCTGACATTTCTCAAGCTTGTCTTACCTGTGCAACATTTGGCCACC 1171
 QY 481 CGAATTTCTTCCCAATCTTTTATCTTGGTGCAGAGGAATGTCCTCAACAGAGCTATA 540
 DB 1172 CGAATTTCTTCCCAATCTTTTATCTTGGTGCAGAGGAATGTCCTCAACAGAGCTATA 1231
 QY 541 CAGCAGAAATGGGATTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCCCTGACTT 600
 DB 1232 CAGCAGAAATGGGATTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCCCTGACTT 1291
 QY 601 ATCCCCAGTCTCAATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAAGAAATTTTG 660
 DB 1292 ATCCCCAGTCTCAATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAAGAAATTTTG 1351
 QY 661 CCGTGTGTTGACAAATAGTATGATTTTATTTGAGAAAGCAAGCTTCAATGATGTGTT 720
 DB 1352 CCGTGTGTTGACAAATAGTATGATTTTATTTGAGAAAGCAAGCTTCAATGATGTGTT 1411
 QY 721 CTAGTCACTGTTAGTGGGATCTCCGCTCCGACCAATCGCTATCGCTCATCATG 780
 DB 1412 CTAGTCACTGTTAGTGGGATCTCCGCTCCGACCAATCGCTATCGCTCATCATG 1471
 QY 781 AAGAGATGACATGTCCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
 DB 1472 AAGAGATGACATGTCCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1531
 QY 841 ATATCTTCAAACTTCAATTTTCTGGGCAACTCTGTGACTATGAGAAAGATTTAGAAC 900
 DB 1532 ATATCTTCAAACTTCAATTTTCTGGGCAACTCTGTGACTATGAGAAAGATTTAGAAC 1591
 QY 901 CAGACTGAGATCAGGGCCCAAGAGCAAACTCAAGCTGTGCACTGTGAAGCCCAAT 960
 DB 1592 CAGACTGAGATCAGGGCCCAAGAGCAAACTCAAGCTGTGCACTGTGAAGCCCAAT 1651
 QY 961 GAACCTGTCCCTGTCTCAGAGAGGTGACAGAAAAAGAGAGCCCTCACTGTCACCC 1020
 DB 1652 GAACCTGTCCCTGTCTCAGAGAGGTGACAGAAAAAGAGAGCCCTCACTGTCACCC 1711

QY 1021 TGTGCCGACTCTGCTACTCTCAGAGGACAGAGCAAAAGGCCCTGTGATCCCGCAGCGTG 1080
 DB 1712 TGTGCCGACTCTCTACTCTCAGAGGACAGAGCAAAAGGCCCTGTGATCCCGCAGCGTG 1771
 QY 1081 CCCAGCGTCCCGCAGCGTGCAGCGCTGCTGTTTGAAGACAGCCCGCTGTATACAGCGCTC 1140
 DB 1772 CCCAGCGTCCCGCAGCGTGCAGCGCTGCTGTTTGAAGACAGCCCGCTGTATACAGCGCTC 1831
 QY 1141 AGTGGCTGCACCTGTCCGACAGCAGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
 DB 1832 AGTGGCTGCACCTGTCCGACAGCAGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1891
 QY 1201 TCTGTGATATCAAAATCAGTTTATATATTCAGCCAGATGGCAGATCTTACATGCTTC 1260
 DB 1892 TCTGTGATATCAAAATCAGTTTATATATTCAGCCAGATGGCAGATCTTACATGCTTC 1951
 QY 1261 TCTCATCAGAAAGTCTTGTGAATACTACAAACCTTCCACTACTGTGATGGAACCAAC 1320
 DB 1952 TCTCATCAGAAAGTCTTGTGAATACTACAAACCTTCCACTACTGTGATGGAACCAAC 2011
 QY 1321 AAGCTATGCCAGTTCTTCCCTGTTTCAAGAACTATCGAGAGACTCCGAAACCACTCT 1380
 DB 2012 AAGCTATGCCAGTTCTTCCCTGTTTCAAGAACTATCGAGAGACTCCGAAACCACTCT 2071
 QY 1381 GATTAAGGAGAGCCAGCATCCCAAGAGCTGCAGACGCGCAGGCTTGACAGCGCAG 1440
 DB 2072 GATTAAGGAGAGCCAGCATCCCAAGAGCTGCAGACGCGCAGGCTTGACAGCGCAG 2131
 QY 1441 AGCAAGCATTTGATTCGTCAGAACAGCAGAGTGGCAACCGCCAGAGGCTCTTTTA 1500
 DB 2132 AGCAAGCATTTGATTCGTCAGAACAGCAGAGTGGCAACCGCCAGAGGCTCTTTTA 2191
 QY 1501 TCTTCACCTGCATGAAAGTGGAGCGTGGAGGACAAATTAACAACAGCTTCTTTTGGC 1560
 DB 2192 TCTTCACCTGCATGAAAGTGGAGCGTGGAGGACAAATTAACAACAGCTTCTTTTGGC 2251
 QY 1561 CTTTCACAGCAGCAGAGACACCTCAGAAAGTCTGAGGCTTGAAGGGCTGGCAC 1620
 DB 2252 CTTTCACAGCAGCAGAGACACCTCAGAAAGTCTGAGGCTTGAAGGGCTGGCAC 2311
 QY 1621 TCGGATATCTTGGCCCCCGCAGACCTCTACCCCTTCCCTGACAGAGCTGTATTGCGC 1680
 DB 2312 TCGGATATCTTGGCCCCCGCAGACCTCTACCCCTTCCCTGACAGAGCTGTATTGCGC 2371
 QY 1681 ACAGAGTCTCAGACTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTCCAGTTACTT 1740
 DB 2372 ACAGAGTCTCAGACTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTCCAGTTACTT 2431
 QY 1741 GCCTAAGCTGCAGCGAGCTGCCACTTGGGAGACCAAGTCTATTCTGGGCGCAGGCGG 1800
 DB 2432 GCCTAAGCTGCAGCGAGCTGCCACTTGGGAGACCAAGTCTATTCTGGGCGCAGGCGG 2491
 QY 1801 CAGAAACCAAGTGCAGAGCTGACTCGCGCGAGCTGCGATGAGAGAGCCCTTTGAA 1860
 DB 2492 CAGAAACCAAGTGCAGAGCTGACTCGCGCGAGCTGCGATGAGAGAGCCCTTTGAA 2551
 QY 1861 AAGCAGTTTAAACGAGAGCTGCCAATGGAATTTGAGAGAGCATATGTCAAGAAC 1920
 DB 2552 AAGCAGTTTAAACGAGAGCTGCCAATGGAATTTGAGAGAGCATATGTCAAGAAC 2611
 QY 1921 AGGTCAACGGGAAGAGCTGGGGAAAGTGGGAGTCACTTCTTCTGGGAGACATGGA 1980
 DB 2612 AGGTCAACGGGAAGAGCTGGGGAAAGTGGGAGTCACTTCTTCTGGGAGACATGGA 2671
 QY 1981 ATCATTTAGGTCCTCC 1995
 DB 2672 ATCATTTAGGTCCTCC 2686

RESULT 8
 US-10-343-357-17
 ; Sequence 17, Application US/10343357
 ; Publication No. US20040058341A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
APPLICANT: VAO, Monique G.; BURFORD, Neil
APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
APPLICANT: GANDHI, Ameeta R.; ARVIZU, Chandra S.
APPLICANT: LEE, Ernestine A.; HAFALIA, April J.A.
APPLICANT: LU, Dzung Aina M.; TRIBOULEY, Catherine M.
APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
APPLICANT: YUE, Henry; WARREN, Bridget A.
APPLICANT: NGUYEN, Daniel B.; CHAMLA, Narinder K.
APPLICANT: KEARNEY, Liam
TITLE OF INVENTION: PROTEIN PHOSPHATASES
FILE REFERENCE: PI-0173 PCT
CURRENT APPLICATION NUMBER: US/10/343,357
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,679
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/223,272
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/224,309
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 60/226,728
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/229,254
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: US 60/231,366
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 17
LENGTH: 3766
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CB1
US-10-343-357-17

Query Match

Best Local Similarity 99.8%; Score 1991.8; DB 13; Length 3766;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCCCATGAGATGATGTTGAATCTCAATTTGTTACTGAGAGGTTGCTGCTGCGAA 60
DB 538 ATGCCCATGAGATGATGTTGAATCTCAATTTGTTACTGAGAGGTTGCTGCTGCGAA 60
QY 61 AGTGGAAACGGAAGAGTGTCTAATGATAGCCGCAATTTGGAATAAATATACATCC 120
DB 598 AGTGGAAACGGAAGAGTGTCTAATGATAGCCGCAATTTGGAATAAATATACATCC 120
QY 121 CACATTTTGAAGCCATTAATATCACTGCTCCAAAGCTTATGAGCAAGGTTGCAACG 180
DB 658 CACATTTTGAAGCCATTAATATCACTGCTCCAAAGCTTATGAGCAAGGTTGCAACG 180
QY 181 GACAAAGTTTAAATACAGAGCTCATCCAGATTCAAGGAAATTAAGTTGACATTGAT 240
DB 718 GACAAAGTTTAAATACAGAGCTCATCCAGATTCAAGGAAATTAAGTTGACATTGAT 240
QY 241 TGCAGTGAAGAGTTGATGTTAAGTCAAGTCCCAAGATGTTGCTGCTGCTTCA 300
DB 778 TGCAGTGAAGAGTTGATGTTAAGTCAAGTCCCAAGATGTTGCTGCTGCTTCA 300
QY 301 GACTGTTTCTCACTGTTACTTGGGTTAACTGGAAGAAGCTTCACTGTTCACTG 360
DB 838 GACTGTTTCTCACTGTTACTTGGGTTAACTGGAAGAAGCTTCACTGTTCACTG 360
QY 361 CTTCAGAGTGGTTGCTGAGTCTCTCGTTGTTCCCTGCGCTCTGGAAGAAATCC 420
DB 898 CTTCAGAGTGGTTGCTGAGTCTCTCGTTGTTCCCTGCGCTCTGGAAGAAATCC 420

QY 421 ACTTAGTCCCTACCTGCAATTTCTAGCCTTGTCTTACCTGTGGCAACATTTGGCCAAAC 480
DB 958 ACTTAGTCCCTACCTGCAATTTCTAGCCTTGTCTTACCTGTGGCAACATTTGGCCAAAC 480
QY 481 CGAATTTCTCCCAATCTTATCTTGGTGGCCAGAGAGTGGCTTCAAGAGAGCTGATG 540
DB 1018 CGAATTTCTCCCAATCTTATCTTGGTGGCCAGAGAGTGGCTTCAAGAGAGCTGATG 540
QY 541 CAGCAAGATGGAGTTGATGTTAAGTGAATCCAGCTATACCTGTCCAAAGCTGACTT 600
DB 1078 CAGCAAGATGGAGTTGATGTTAAGTGAATCCAGCTATACCTGTCCAAAGCTGACTT 600
QY 601 ATCCCGAGTCTCAATTTCTGCGTGTCTGTGAATGACAGCTTTGGAAGAAATTTTG 660
DB 1138 ATCCCGAGTCTCAATTTCTGCGTGTCTGTGAATGACAGCTTTGGAAGAAATTTTG 660
QY 661 CCGTGTGGAGCAATAGTGAATTTCTTGAAGAAAGCAAGCTTCAATGATGTT 720
DB 1198 CCGTGTGGAGCAATAGTGAATTTCTTGAAGAAAGCAAGCTTCAATGATGTT 720
QY 721 CTAGTGAATGTTAGTGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 780
DB 1258 CTAGTGAATGTTAGTGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 780
QY 781 AAGAGATGAGCAATGCTTATGATGAAGCTTACAGATTTGTAAGAAAGAAAGCTACT 840
DB 1318 AAGAGATGAGCAATGCTTATGATGAAGCTTACAGATTTGTAAGAAAGAAAGCTACT 840
QY 841 ATATCTCCAAATCTTCAATTTTGGGCAACTCTGCACTATGAGAAAGATTAAGAAC 900
DB 1378 ATATCTCCAAATCTTCAATTTTGGGCAACTCTGCACTATGAGAAAGATTAAGAAC 900
QY 901 CAGACTGAGCATCAGGCGCAAGAGCAACTGAGCTGCTGCACTGAGAAAGCCAAAT 960
DB 1438 CAGACTGAGCATCAGGCGCAAGAGCAACTGAGCTGCTGCACTGAGAAAGCCAAAT 960
QY 961 GAACCTGCTCCGCTGTCTCAAGAGGTGACAGAAAGGAGACGCTTCACTGCAACC 1020
DB 1498 GAACCTGCTCCGCTGTCTCAAGAGGTGACAGAAAGGAGACGCTTCACTGCAACC 1020
QY 1021 TGTGCGCACTGTGTTACTCAGAGGTCGACAGAAAGCCGCTGATCCGCAAGCTG 1080
DB 1558 TGTGCGCACTGTGTTACTCAGAGGTCGACAGAAAGCCGCTGATCCGCAAGCTG 1080
QY 1081 CCGAGCGTCCGAGCGCTGACGCGTGTGTAGAGACAGCGCGCTGTACAGGCGCTC 1140
DB 1618 CCGAGCGTCCGAGCGCTGACGCGTGTGTAGAGACAGCGCGCTGTACAGGCGCTC 1140
QY 1141 AGTGGCTGACCTGTCCGACAGAGGCTGGAAGACAGATTAAGCTCAAGCTTCC 1200
DB 1678 AGTGGCTGACCTGTCCGACAGAGGCTGGAAGACAGATTAAGCTCAAGCTTCC 1200
QY 1201 TCTCTGATATCAATCAAGTTTCAATTTACCGACAGATGAGCATCTTACATGCTTC 1260
DB 1738 TCTCTGATATCAATCAAGTTTCAATTTACCGACAGATGAGCATCTTACATGCTTC 1260
QY 1261 TCTCTATCGAAGATGCTTTGGAATATCAAAACCTTCCATCTGTGATGAGCAAC 1320
DB 1798 TCTCTATCGAAGATGCTTTGGAATATCAAAACCTTCCATCTGTGATGAGCAAC 1320
QY 1321 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATGGAAGAGACTCCGAACCAAGTCT 1380
DB 1858 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATGGAAGAGACTCCGAACCAAGTCT 1380
QY 1381 GATTAAGGAGAGCCAGATCTCCCAAGAGCTGACAGCCGCAAGCTTCAAGACGCGAG 1440
DB 1918 GATTAAGGAGAGCCAGATCTCCCAAGAGCTGACAGCCGCAAGCTTCAAGACGCGAG 1440
QY 1441 AGCAAGGATTTGATTTGCTGCAAGACAGAGCTGCAAGCTTCAAGACGCGAG 1500
DB 1978 AGCAAGGATTTGATTTGCTGCAAGACAGAGCTGCAAGCTTCAAGACGCGAG 1500
QY 1501 TCTCACTGATGAAGTGGAGCTGAGAGCAATTAACACACAGCTTCTTTTCGGC 1560

Db 2038 TCTCCATGATGGAAGTGGAGGTGAGAGACATTAACACACAGCTTCTTTGGC 2097
Qy 1561 CTTTCCACCCGACGACGACCTCAAGAGTCTGCTGGCTTGAAGGCTGGCAC 1620
Db 2098 CTTTCCACCCGACGACGACCTCAAGAGTCTGCTGGCTTGAAGGCTGGCAC 2157
Qy 1621 TCGGATCTTGGGCCCCGAGCCTTAACCCCTTCCCTGACCGAGCTGGTATTTGGC 1680
Db 2158 TCGGATATCTTGGGCCCCGAGCCTTAACCCCTTCCCTGACCGAGCTGGTATTTGGC 2217
Qy 1681 ACAGAGCTCTACACTTCTACTCTGCTGACCCATCTACGAGGAGCTGCACTTCT 1740
Db 2218 ACAGAGCTCTACACTTCTACTCTGCTGACCCATCTACGAGGAGCTGCACTTCT 2277
Qy 1741 GCCTACAGCTGACGACGCTGCCCACTTGGCGAGACCAAGTCTATTTCTGGCAGCGG 1800
Db 2278 GCCTACAGCTGACGACGCTGCCCACTTGGCGAGACCAAGTCTATTTCTGGCAGCGG 2337
Qy 1801 CAGAAGCCAGTGAAGAGCTGACCTGGCGGCGAGCTGGAGTGAAGAGAGCCCTTGA 1860
Db 2338 CAGAAGCCAGTGAAGAGCTGACCTGGCGGCGAGCTGGAGTGAAGAGAGCCCTTGA 2397
Qy 1861 AACGAGTTTAAAGCAGAGAGCTGCAATGAAATTTGAGAGAGCATCATGTCAGAGAC 1920
Db 2398 AACGAGTTTAAAGCAGAGAGCTGCAATGAAATTTGAGAGAGCATCATGTCAGAGAC 2457
Qy 1921 AGGTACGAGAGAGCTGGGAGAAAGTGGGCACTGCTGCTTTGGCGAGCAGTGA 1980
Db 2458 AGGTACGAGAGAGCTGGGAGAAAGTGGGCACTGCTGCTTTGGCGAGCAGTGA 2517
Qy 1981 ATCATTTAGGCTGCC 1995
Db 2518 ATCATTTAGGCTGCC 2532

RESULT 9

US-10-648-593-115
; Sequence 115, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 115
; LENGTH: 4790
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-648-593-115

Query Match 99.8%; Score 1991.8; DB 17; Length 4790;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATTGGAACTCAATTTGTTACTGAGAGGTTGGTGGCTTGGCTGGA 60
Db 184 ATGGCCCATGAGATGATTGGAACTCAATTTGTTACTGAGAGGTTGGTGGCTTGGCTGGA 243
Qy 61 AGTGAAGCGGAAAGTGTCTTAATTGATAGCCGGCATTGTTGGAATAACAATATCC 120
Db 244 AGTGAAGCGGAAAGTGTCTTAATTGATAGCCGGCATTGTTGGAATAACAATATCC 303
Qy 121 CACATTTGGAAGCATTAAATATCACTGCTCAAGGTTATGAAGCAGAGTTGCAACAG 180
Db 304 CACATTTGGAAGCATTAAATATCACTGCTCAAGGTTATGAAGCAGAGTTGCAACAG 363

Qy 181 GACAAAGTTTAAATTCAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 240
Db 364 GACAAAGTTTAAATTCAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 423
Qy 241 TGCAGTCAGAGGTTGATGTTACGATCAAGTCTCCAGAGATGTCCTCTCTTCA 300
Db 424 TGCAGTCAGAGGTTGATGTTACGATCAAGTCTCCAGAGATGTCCTCTCTTCA 483
Qy 301 GACTGTTTCTCACTGTACTTCTGGGTAACTGGAAGAGCTTCAACTGTTCACCTG 360
Db 484 GACTGTTTCTCACTGTACTTCTGGGTAACTGGAAGAGCTTCAACTGTTCACCTG 543
Qy 361 CTTCGAGGTTGGTTGCTGAGTTCTCGTGGTTTCCCTGGAGCTGTGAGAGAAATCC 420
Db 544 CTTCGAGGTTGGTTGCTGAGTTCTCGTGGTTTCCCTGGAGCTGTGAGAGAAATCC 603
Qy 421 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGGCTTACTGTTGCCAATTTGGCCAAC 480
Db 604 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGGCTTACTGTTGCCAATTTGGCCAAC 663
Qy 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGAGAGATGTCCTCAACAGAGCTGATG 540
Db 664 CGAATTTCTCCCAATCTTTATCTTGGCTGCGAGAGATGTCCTCAACAGAGCTGATG 723
Qy 541 CAGCAGAAATGGGATTTGTTATGTTAAATGCCAGCTATTCCTGTCAAAGCCTGACTT 600
Db 724 CAGCAGAAATGGGATTTGTTATGTTAAATGCCAGCTATTCCTGTCAAAGCCTGACTT 783
Qy 601 ATCCCGAGCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 784 ATCCCGAGCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843
Qy 661 CCGTGTGGAACAATTCAGTATGATTTCATTGAGAAAGCAAAAGCCTTCAATGATGTT 720
Db 844 CCGTGTGGAACAATTCAGTATGATTTCATTGAGAAAGCAAAAGCCTTCAATGATGTT 903
Qy 721 CTAGTACAGCTTTTACTGGGATCTCCGCTCCGACCAATGCTATGCTATCATCATG 780
Db 904 CTAGTACAGCTTTTACTGGGATCTCCGCTCCGACCAATGCTATGCTATCATCATG 963
Qy 781 AAGAGATGAGACATGCTTTTAAATGAGAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 840
Db 964 AAGAGATGAGACATGCTTTTAAATGAGAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 1023
Qy 841 AATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
Db 1024 AATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 1083
Qy 901 CAGACTGAGCATCAGAGGCAAGAGCAAAAGCTCAAGCTGTCACCTGGAGAGCCAAAT 960
Db 1084 CAGACTGAGCATCAGAGGCAAGAGCAAAAGCTCAAGCTGTCACCTGGAGAGCCAAAT 1143
Qy 961 GAACCTGTCTGCTGCTCTCAGAGGTTGAGAGAGAAAGCGAGAGAGGCTTCACTCAACC 1020
Db 1144 GAACCTGTCTGCTGCTCTCAGAGGTTGAGAGAGAGAAAGCGAGAGGCTTCACTCAACC 1203
Qy 1021 TGTGCGACTCTGCTACTCAGAGGAGAGAGCAAAAGGCTGTCATCCGACAGGTG 1080
Db 1204 TGTGCGACTCTGCTACTCAGAGGAGAGAGCAAAAGGCTGTCATCCGACAGGTG 1263
Qy 1081 CCCAGGTCGCCAGGCTGAGAGCGGTGCTGTTAGAGAGACAGCCGCTGGTGAAGGAGCTC 1140
Db 1264 CCCAGGTCGCCAGGCTGAGAGCGGTGCTGTTAGAGAGACAGCCGCTGGTGAAGGAGCTC 1323
Qy 1141 AGTGGCTGACCTGTCTCCGAGAGCAGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1324 AGTGGCTGACCTGTCTCCGAGAGCAGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1383
Qy 1201 TCTCTGATATCAATCAAGTTTCAATTTAGAGCAGATGAGCAGATCTTATCATGCTTC 1260
Db 1384 TCTCTGATATCAATCAAGTTTCAATTTAGAGCAGATGAGCAGATCTTATCATGCTTC 1443

QY 1261 TCCTCATCAGAGAGTGTGGATGTAATCTAACCTTCCACTACTGTGATGGAGCAAC 1320
DB 1444 TCCTCATCAGAGAGTGTGGATGTAATCTAACCTTCCACTACTGTGATGGAGCAAC 1503
QY 1321 AAGCTATGCAAGTTCCTCCCTGTTCCAGAACTATGGAGCAAGCTCCCGAAACCAAGTCT 1380
DB 1504 AAGCTATGCAAGTTCCTCCCTGTTCCAGAACTATGGAGCAAGCTCCCGAAACCAAGTCT 1563
QY 1381 GATTAAGAGAGAGCAAGTTCCTCCCGAAAGCTGCAAGACCGCAGGCTTTCAGACCAAG 1440
DB 1564 GATTAAGAGAGAGCAAGTTCCTCCCGAAAGCTGCAAGACCGCAGGCTTTCAGACCAAG 1623
QY 1441 AGCAAGCAATTCATTCGCTGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1500
DB 1624 AGCAAGCAATTCATTCGCTGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1683
QY 1501 TCCTCATCAGAGAGTGTGGATGTAATCTAACCTTCCACTACTGTGATGGAGCAAC 1560
DB 1684 TCCTCATCAGAGAGTGTGGATGTAATCTAACCTTCCACTACTGTGATGGAGCAAC 1743
QY 1561 CTTCACACAGCCAGCAAGCTTCAAGAGTCTGAGCAAGCAAGCAAGCAAGCAAGCAAG 1620
DB 1744 CTTCACACAGCCAGCAAGCTTCAAGAGTCTGAGCAAGCAAGCAAGCAAGCAAGCAAG 1803
QY 1621 TCAGATATCTGGGCCCCCAGACCTTACCTTCCCTGACCAAGCAAGCAAGCAAGCAAG 1680
DB 1804 TCAGATATCTGGGCCCCCAGACCTTACCTTCCCTGACCAAGCAAGCAAGCAAGCAAG 1863
QY 1681 ACAGAGCTCTCAACTTCTACTCTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1740
DB 1864 ACAGAGCTCTCAACTTCTACTCTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1923
QY 1741 GCTTACAGCTGAGCAAGCAAGCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1800
DB 1924 GCTTACAGCTGAGCAAGCAAGCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1983
QY 1801 CAGAAGCCCAAGTGAAGAGCTGACTCGCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAG 1860
DB 1984 CAGAAGCCCAAGTGAAGAGCTGACTCGCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAG 2043
QY 1861 AAGCAGTTTAAACGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
DB 2044 AAGCAGTTTAAACGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2103
QY 1921 AGGTCAAGGAGAGAGTGAAG 1980
DB 2104 AGGTCAAGGAGAGAGTGAAG 2163
QY 1981 ATCATTGAGGTCTCC 1995
DB 2164 ATCATTGAGGTCTCC 2178

RESULT 10
US-10-168-506-2
Sequence 2, Application US/10168506
Publication No. US20040053229A1
GENERAL INFORMATION:
APPLICANT: PLOMAN, GREGORY D.
APPLICANT: MARTINEZ, RICARDO
APPLICANT: WHITE, DAVID
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
APPLICANT: HILL, RON
TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
FILE REFERENCE: 038602/1351
CURRENT APPLICATION NUMBER: US/10/168, 506
PRIOR FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: PCT/US00/34736
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2
LENGTH: 2732
TYPE: DNA
ORGANISM: Homo sapiens
US-10-168-506-2

Query Match 99.8%; Score 1990.2; DB 13; Length 2732;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAATCAATTTGTTACTGAGAGGTGGTGGCTGTGCTGGAA 60
DB 538 ATGGCCCATGAGATGATGGAATCAATTTGTTACTGAGAGGTGGTGGCTGTGCTGGAA 597
QY 61 AGTGAAGGAAAAAGTGGCTTAATGATAGCCGCGCAATTTGTGATATACATATCATCC 120
DB 598 AGTGAAGGAAAAAGTGGCTTAATGATAGCCGCGCAATTTGTGATATACATATCATCC 657
QY 121 CACATTTTGAAGCATTAAATATCAATGCTCCCAAGCTTATGAGAGAGAGAGAGAGAG 180
DB 658 CACATTTTGAAGCATTAAATATCAATGCTCCCAAGCTTATGAGAGAGAGAGAGAGAG 717
QY 181 GACAAAGTGTAAATTCAGAGCTCAATCCAGCAATTCAGGAAACATTAAGTTGACATTGAT 240
DB 718 GACAAAGTGTAAATTCAGAGCTCAATCCAGCAATTCAGGAAACATTAAGTTGACATTGAT 777
QY 241 TGCAGTCAGAGGTTTGAATTTAGATCAAGATCCCAAGATGTTGCTCTCTCTTCA 300
DB 778 TGCAGTCAGAGGTTTGAATTTAGATCAAGATCCCAAGATGTTGCTCTCTCTTCA 837
QY 301 GACTGTTTCTCAGTACTTCTGAGTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 838 GACTGTTTCTCAGTACTTCTGAGTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897
QY 361 CTTCAGAGTGGGTTTGGTGAATTTCTGTTGTTTCCCTGAGCTCTGTGAGAGAGAGAGAG 420
DB 898 CTTCAGAGTGGGTTTGGTGAATTTCTGTTGTTTCCCTGAGCTCTGTGAGAGAGAGAGAG 957
QY 421 ACTTGAATCCCTACCTGATTTCTCAGCTTTCCTGCTTACCTGTTCCCAATTTGGGCAAC 480
DB 958 ACTTGAATCCCTACCTGATTTCTCAGCTTTCCTGCTTACCTGTTCCCAATTTGGGCAAC 1017
QY 481 CGAATTCCTCCCAATCTTTATCTTGGCTGCAAGAGATGCTCTCAACAAGAGAGAGAGAG 540
DB 1018 CGAATTCCTCCCAATCTTTATCTTGGCTGCAAGAGATGCTCTCAACAAGAGAGAGAGAG 1077
QY 541 CAGCAGATGGAGATGTTATGTTAAATGCACTTACCTGCTCCCAAGCTGACTTT 600
DB 1078 CAGCAGATGGAGATGTTATGTTAAATGCACTTACCTGCTCCCAAGCTGACTTT 1137
QY 601 ATCCCGAGTCTGATTCCTGGGCTGCTGCTGTAATGACAGCTTTTGAGAAAAATTTTG 660
DB 1138 ATCCCGAGTCTGATTCCTGGGCTGCTGCTGTAATGACAGCTTTTGAGAAAAATTTTG 1197
QY 661 CGGTGTTGAACAATCAGTATGATTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 1198 CGGTGTTGAACAATCAGTATGATTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
QY 721 CTAGTGAATGTTTGGTGGAGTCTCCCGCTCCGCAACATGCTTATGCTATGATG 780
DB 1258 CTAGTGAATGTTTGGTGGAGTCTCCCGCTCCGCAACATGCTTATGCTATGATG 1317
QY 781 AAGAGATGAGATGCTTTTGAATGAGCTTACAGATTTGTGAAGAGAGAGAGAGAGAGAG 840
DB 1318 AAGAGATGAGATGCTTTTGAATGAGCTTACAGATTTGTGAAGAGAGAGAGAGAGAGAG 1377
QY 841 ATATCTCCAAATCTCAATTTTGGGCGCAATCTGAGCTATGAGAGAGAGAGAGAGAGAG 900
DB 1378 ATATCTCCAAATCTCAATTTTGGGCGCAATCTGAGCTATGAGAGAGAGAGAGAGAGAG 1437
QY 901 CAGAGTGAAGATCAAGGCTCAAGAGCAATCAAGCTGCTGCACTGGAGAGAGAGAGAGAG 960
DB 1438 CAGAGTGAAGATCAAGGCTCAAGAGCAATCAAGCTGCTGCACTGGAGAGAGAGAGAGAG 1497

OY	961	GAACCTGTCCTCCGTGCTCTCAGAGGGGTGAGACAGAAACACAGACGCCCTCAGTCCACCC	102
Db	1498	GAACCTGTCCTCCGTGCTCTCAGAGGGGTGAGACAGAAACACAGACGCCCTCAGTCCACCC	1557
OY	1021	TGTGCGCACTCTGCTACTCTCAGAGGACAGAGACAAAGGCCCTGTCATCCCGCAGCGTG	1080
Db	1558	TGTGCGCACTCTGCTACTCTCAGAGGACAGAGACAAAGGCCCTGTCATCCCGCAGCGTG	1617
OY	1081	CCGACGGGCCACAGGTGACAGCCGTGCTGTGAAGGACAGCCCGCTGTGACAGCGCTC	1140
Db	1618	CCGACGGGCCACAGGTGACAGCCGTGCTGTGAAGGACAGCCCGCTGTGACAGCGCTC	1677
OY	1141	AGTGGGCTGCACCTGTCCGACAGACAGCGCTGAGAGACAGCAATTAAGCTCAAGCGTTCCTTC	1200
Db	1678	AGTGGGCTGCACCTGTCCGACAGACAGCGCTGAGAGACAGCAATTAAGCTCAAGCGTTCCTTC	1737
OY	1201	TCTCTGGAATATCAAAATCAAGTTTCATATTCAGCCAGCATGCGAGCATCTTACATGCGCTTC	1266
Db	1738	TCTCTGGAATATCAAAATCAAGTTTCATATTCAGCCAGCATGCGAGCATCTTACATGCGCTTC	1797
OY	1261	TCCTCATCAGAAAGATGCTTGGAAATACACAACTTCACATCTCTGATGGGACCAAC	1302
Db	1798	TCCTCATCAGAAAGATGCTTGGAAATACACAACTTCACATCTCTGATGGGACCAAC	1857
OY	1321	AAGCTATGCCAGTTCTCCCTGTTTCAGGAATATCGAGACATTCGAGACATCCCGAAACAGTCT	1380
Db	1858	AAGCTATGCCAGTTCTCCCTGTTTCAGGAATATTCGAGACATTCGAGACATCCCGAAACAGTCT	1917
OY	1381	GATTAAGGAGGAAGCCAGCATCCCCAGAAAGCTGCGACACCGCCAGGCTTCAGACAGCCAG	1440
Db	1918	GATTAAGGAGGAAGCCAGCATCCCCAGAAAGCTGCGACACCGCCAGGCTTCAGACAGCCAG	1977
OY	1441	AGCAAGCATTTGCAATTCGCTGAGAACACAGACAGTGGCACCGCCAGAGGTCCCTTTTA	1500
Db	1978	AGCAAGCATTTGCAATTCGCTGAGAACACAGACAGTGGCACCGCCAGAGGTCCCTTTTA	2037
OY	1501	TCTCCACCTGCATCCGAATGGGAGGCTGGAGACAAATTACACACAGCTTCCTTTTGCGC	1566
Db	2038	TCTCCACCTGCATCCGAATGGGAGGCTGGAGACAAATTACACACAGCTTCCTTTTGCGC	2097
OY	1561	CTTTCCACACAGACAGACACCTCAGCAAGATCTGTGGCTTGAAGGGCTGGAC	1620
Db	2098	CTTTCCACACAGACAGACACCTCAGCAAGATCTGTGGCTTGAAGGGCTGGAC	2157
OY	1621	TCGGATATCTTGCCCCCCAGACCTCTACCCCTTCCTCTGACACAGACGTGTGATTTTGCC	1680
Db	2158	TCGGATATCTTGCCCCCCAGACCTCTACCCCTTCCTCTGACACAGACGTGTGATTTTGCC	2217
OY	1681	ACAGAGTCTCTCAACCTTACTCTGTGCTCAAGCATCTTAGGAGGACGTGCCAGTTACTCT	1740
Db	2218	ACAGAGTCTCTCAACCTTACTCTGTGCTCAAGCATCTTAGGAGGACGTGCCAGTTACTCT	2277
OY	1741	GCTTACAGCTGCAGCGCAGCTGCCCATCTTGCGGAGACCAAGTCTTATCTGTGCGGAGCGG	1800
Db	2278	GCTTACAGCTGCAGCGCAGCTGCCCATCTTGCGGAGACCAAGTCTTATCTGTGCGGAGCGG	2337
OY	1801	CAGAAGCCAAGTACAGAGCTGACTCGCGGCGGAGCTGGCATGAGAGAGGCCCTTTGAA	1860
Db	2338	CAGAAGCCAAGTACAGAGCTGACTCGCGGCGGAGCTGGCATGAGAGAGGCCCTTTGAA	2397
OY	1861	AAGCAGTTTAACCGAGAGAGCTGCGCAATTTGGAGAGACATCATCTCAGAGAAC	1920
Db	2398	AAGCAGTTTAACCGAGAGAGCTGCGCAATTTGGAGAGACATCATCTCAGAGAAC	2457
OY	1921	AGGTCAACGGGAAGAGCTGGGGAAGATGGGACATGACTTACTTTTGGGACAGCATGGAA	1980
Db	2458	AGGTCAACGGGAAGAGCTGGGGAAGATGGGACATGACTTACTTTTGGGACAGCATGGAA	2517
OY	1981	ATCATTAGAGTCTCC	1995
Db	2518	ATCATTAGAGTCTCC	2532

```

RESULT 11
US-10-094-749-673
; Sequence 673, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: MAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SERI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094, 749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 673
; LENGTH: 2102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-673

Query Match      99.7%; Score 1988.6; DB 16; Length 2102;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY      1  ATGSCCCATGAGATGATGGAACCTCAATTGTACTGAGAGGTTGGTCTGCTGCGAA 60
Db      56  ATGSCCCATGAGATGATGGAACCTCAANTTTTACTGAGAGGTTGGTCTGCTGCGAA 115

OY      61  AGTGAACGGAAGAAAGTCTGCTTAATTGATGCGCGCCATTGTGGAATACATACATCC 120
Db      116  AGTGAACGGAAGAAAGTCTGCTTAATTGATGATGCCGCCATTGTGGAATACATACATCC 175

OY      121  CACATTTTGGAAAGCCATTATATCACTGCTCCAAAGCTTATGAAGCGAAAGTTGCAACAG 180
Db      176  CACATTTTGGAAAGCCATTATATCACTGCTCCAAAGCTTATGAAGCGAAAGTTGCAACAG 235

OY      181  GACAAAGTGTTAATTACAGAGCTCATCCAGACATTCAGAGAAACATTAAGTTGACATTGAT 240
Db      236  GACAAAGTGTTAATTACAGAGCTCATCCAGACATTCAGAGAAACATTAAGTTGACATTGAT 295

OY      241  TGCAGTCAGAAAGTTGATTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTTCA 300
Db      296  TGCAGTCAGAAAGTTGATTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTTCA 355

OY      301  GACTGTTTTCTCACTGTACTTCTGGGTAAACTGAGAAAGAGCTTCAACTCTGTTCACTG 360
Db      356  GACTGTTTTCTCACTGTACTTCTGGGTAAACTGAGAAAGAGCTTCAACTCTGTTCACTG 415

OY      361  CTTGCAAGTGGGTTTGGCTGAGTTCTCTCGATGTTGTTCCCTGGCCCTCTGGAAGAAAAATCC 420
Db      416  CTTGCAAGTGGGTTTGGCTGAGTTCTCTCGATGTTGTTCCCTGGCCCTCTGGAAGAAAAATCC 475

OY      421  ACTCTAATCCCTACATGCAATTTCTACGCTTGCTTAACTGTTGCCAACATTTGGGCAACC 480

```


Db 476 ACTGAGTCCCTACCTGCAATTTCTCAGGCTGCTTACCTGTGCAACATTTGGGCGAAC 535
Qy 481 CGAATTCCTCCCAATTTTATTTGCTGCTGCGCAGAGATGTCCTCAACAGAGAGCTGATA 540
Db 536 CTATATCTTCCCAATCTTATCTTGTGCTGCCAGAGATGTCCTCAACAGAGAGCTGATA 540
Qy 541 CAGCAGAAATGGGATTTGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 595
Db 596 CAGCAGAAATGGGATTTGTTATGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 600
Qy 601 ATCCCGAGTCCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
Db 656 ATCCCGAGTCCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 661 CCGTGTGTTGCAAAATCAGTATGATTTTCAATTTGAGAAAGCAAAAGCTCCCAATGATG 715
Db 716 CCGTGTGTTGCAAAATCAGTATGATTTTCAATTTGAGAAAGCAAAAGCTCCCAATGATG 720
Qy 721 CTAGTGACCTGTTTATGCTGGATCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 775
Db 776 CTAGTGACCTGTTTATGCTGGATCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 780
Qy 781 AAGAGATGACATGCTTTAGATGAGAGCTTACAGATTTGAGAAAGCAAAAGCTTACT 835
Db 836 AAGAGATGACATGCTTTAGATGAGAGCTTACAGATTTGAGAAAGCAAAAGCTTACT 840
Qy 841 ATATCTCCAAATCTTCAATTTTCTGGCCCACTCTCTGCACTATGAGAAAGATTTAAGAC 895
Db 896 ATATCTCCAAATCTTCAATTTTCTGGCCCACTCTCTGCACTATGAGAAAGATTTAAGAC 900
Qy 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTCTGAGAGAGCAAA 955
Db 956 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTCTGAGAGAGCAAA 960
Qy 961 GAACCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015
Db 1016 GAACCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1021 TGTGCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075
Db 1076 TGTGCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy 1081 CCCAGCGTCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1135
Db 1136 CCCAGCGTCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Qy 1141 AGTGGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1195
Db 1196 AGTGGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy 1201 TCTCTGATATCAATCACTTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 1255
Db 1256 TCTCTGATATCAATCACTTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 1260
Qy 1261 TCTCTGATGAGATGCTTTGATATCTAACAACCTTCACTTCTGATGAGAGCAAC 1315
Db 1316 TCTCTATCAAGAAATGCTTTGAAATCTAACAACCTTCACTTCTGATGAGAGCAAC 1320
Qy 1321 AAGCTATGCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1375
Db 1376 AAGCTATGCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Qy 1381 GATAAGAGAAAGCAGATCCCAAGAGCTGCAAGCGGCTTCAACAGCCAG 1435
Db 1436 GATAAGAGAAAGCAGATCCCAAGAGCTGCAAGCGGCTTCAACAGCCAG 1440
Qy 1441 AGCAAGCGATTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1495
Db 1496 AGCAAGCGATTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Qy 1501 TCTCACTGATGAGAGTGGAGCGTGGAGCAATTTCAACAGGCTTCTTTTGGC 1555
Db 1556 TCTCACTGATGAGAGTGGAGCGTGGAGCAATTTCAACAGGCTTCTTTTGGC 1560

Db 1556 TCTCACTGATGAGAGTGGAGCGTGGAGCAATTTCAACAGGCTTCTTTTGGC 1615
Qy 1561 CTTTCCACAGCCAGAGAGCACTTCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1615
Db 1616 CTTTCCACAGCCAGAGAGCACTTCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Qy 1621 TCGGATATCTTGGCCCCCAGAGCCCTTACCCCTTCCCTGACACAGAGCTGATTTTGGC 1675
Db 1676 TCGGATATCTTGGCCCCCAGAGCCCTTACCCCTTCCCTGACACAGAGCTGATTTTGGC 1680
Qy 1681 ACAGAGTCCCTCACTTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1735
Db 1736 ACAGAGTCCCTCACTTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Qy 1741 GCTTACAGCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795
Db 1796 GCTTACAGCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Qy 1801 CAGAAAGCAAGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1855
Db 1856 CAGAAAGCAAGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Qy 1861 AAGGATTTAAGCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1915
Db 1916 AAGGATTTAAGCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Qy 1921 AGGTCAAGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1975
Db 1976 AGGTCAAGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Qy 1981 ATCATTTAGCTTCC 1995
Db 2036 ATCATTTAGCTTCC 2050

RESULT 12
US-10-296-115-520
/ Sequence 520, Application US/10296115
/ Publication No. US20040053248A1
/ GENERAL INFORMATION:
/ APPLICANT: Hysq Inc
/ TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
/ FILE REFERENCE: 784PCT
/ CURRENT APPLICATION NUMBER: US/10/296, 115
/ PRIOR FILING DATE: 2002-11-18
/ PRIOR APPLICATION NUMBER: US09/488, 725
/ PRIOR FILING DATE: 2000-01-21
/ PRIOR APPLICATION NUMBER: US09/552, 317
/ NUMBER OF SEQ ID NOS: 1478
/ SEQ ID NO 520
/ LENGTH: 2966
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(2966)
/ OTHER INFORMATION: n = a, c, t or g
US-10-296-115-520

Query Match 99.7%; Score 1988.6; DB 13; Length 2966;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATTTGAAGTCAATTTGTTACTGAGAGGTTGGTCTGCTGGA 60
Db 23 ATGGCCCATGAGATGATTTGAAGTCAATTTGTTACTGAGAGGTTGGTCTGCTGGA 82
Qy 61 AGTGAAGGAGAAAAAGTGTCTAATTTGAATGATCCGCGCATTTTGTGAATACATATATCC 120
Db 83 AGTGAAGGAGAAAAAGTGTCTAATTTGAATGATCCGCGCATTTTGTGAATACATATATCC 142
Qy 121 CACATTTGGAAGCCATTAATATCAACTGCTCCAGCTTTATGAAGCGAAGTTGCAACAG 180


```

Db      143 CACATTTGGAGGCAATTAATCACTGCTCCAAAGCTTATGAGAGGAAAGTTGCAACAG 202
Qy      181 GACAAAGTTTAAATTAACAGAGCTCATCCAGATTACAGAAACATAAGTTGACATTGAT 240
Db      203 GACAAAGTTTAAATTAACAGAGCTCATCCAGATTACAGAAACATAAGTTGACATTGAT 262
Qy      241 TGCAGTGAAGAGTTGATGTTAAGATCAAAAGCTCCAAAGTTGATGCTCTCTCTCA 300
Db      263 TGCAGTGAAGAGTTGATGTTAAGATCAAAAGCTCCAAAGTTGATGCTCTCTCTCA 322
Qy      301 GACTGTTTTCTCACTGTAATCTTGAGTAACTGAGAAAGAGCTTCAACTCTGCTCACTG 360
Db      323 GACTGTTTTCTCACTGTAATCTTGAGTAACTGAGAAAGAGCTTCAACTCTGCTCACTG 382
Qy      361 CTTCGAGAGTGGGTTTGCTGAGATTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
Db      383 CTTCGAGAGTGGGTTTGCTGAGATTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 442
Qy      421 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGTCTTACCTGTGCAATTTGGGCAACC 480
Db      443 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGTCTTACCTGTGCAATTTGGGCAACC 502
Qy      481 CGAATCTTCCCAATCTTTATCTTGCTGCGCAGAGATGTCCTCAACAAGAGCTGATA 540
Db      503 CGAATCTTCCCAATCTTTATCTTGCTGCGCAGAGATGTCCTCAACAAGAGCTGATA 562
Qy      541 CAGCAGAAATGGGATGGTATGTTAAATGCGCAGCTATACCTGTCCAAAGCTGACTTT 600
Db      563 CAGCAGAAATGGGATGGTATGTTAAATGCGCAGCTATACCTGTCCAAAGCTGACTTT 622
Qy      601 ATCCCGAGTCTATTTCTGCGGTGCTGTAATGACAGCTTTGTGAAAAATTTTG 660
Db      623 ATCCCGAGTCTATTTCTGCGGTGCTGTAATGACAGCTTTGTGAAAAATTTTG 682
Qy      661 CCGTGTGTTGCAAAATCAGTAGATTTTCATTGAGAAACAAAGCTTCAATGATGTT 720
Db      683 CCGTGTGTTGCAAAATCAGTAGATTTTCATTGAGAAACAAAGCTTCAATGATGTT 742
Qy      721 CTATGTCAGCTGTTTAACTGGGATCTCCGCTCGGCACCAATCGGTATGCTTACATCAG 780
Db      743 CTATGTCAGCTGTTTAACTGGGATCTCCGCTCGGCACCAATCGGTATGCTTACATCAG 802
Qy      781 AAGAAGATGAGCAATGCTTTAATGAGAGCTTACAGATTTGTGAAAGAAAGACCTACT 840
Db      803 AAGAAGATGAGCAATGCTTTAATGAGAGCTTACAGATTTGTGAAAGAAAGACCTACT 862
Qy      841 ATATCTCCAATCTTCAATTTTCTGGGCAACTCTGTGACTATGAGAAAGATTTAAGAAC 900
Db      863 ATATCTCCAATCTTCAATTTTCTGGGCAACTCTGTGACTATGAGAAAGATTTAAGAAC 922
Qy      901 CAGACTGAGCAATCAAGGAGCAAAAGAGCAAACTCAAGCTGTGCACTGAGAAAGCAAT 960
Db      923 CAGACTGAGCAATCAAGGAGCAAAAGAGCAAACTCAAGCTGTGCACTGAGAAAGCAAT 982
Qy      961 GAACCTGTCCCTGCTGTCTGAGAGGTTGACAAAAAGCGCAAGCGCTCTCATCTCCACC 1020
Db      983 GAACCTGTCCCTGCTGTCTGAGAGGTTGACAAAAAGCGCAAGCGCTCTCATCTCCACC 1042
Qy      1021 TGTGCGAATCTGTCTGACTCTGAGAGGAGAGCAAAAGCGCGTGCATCCCGCAGAGCTG 1080
Db      1043 TGTGCGAATCTGTCTGACTCTGAGAGGAGAGCAAAAGCGCGTGCATCCCGCAGAGCTG 1102
Qy      1081 CCCAGGCTGCCAGCGTGCAGCGGTGCTGTATGAGAGCAGCCGCTGTATCAGGCGCTC 1140
Db      1103 CCCAGGCTGCCAGCGTGCAGCGGTGCTGTATGAGAGCAGCCGCTGTATCAGGCGCTC 1162
Qy      1141 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCGTCTTC 1200
Db      1163 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCGTCTTC 1222
Qy      1201 TCTCTGATATCAAAATCAGTTTCAATATTCAGCCAGCATGAGCATCTTACATGCTTC 1260

```

```

Db      1223 TCTCTGATATCAAAATCAGTTTCAATATTCAGCCAGCATGAGCATCTTACATGCTTC 1282
Qy      1261 TCTCTCATCAGAAATGCTTTGGAATATCTAACAACTTCCACTCTCTGATGAGGACCAAC 1320
Db      1283 TCTCTCATCAGAAATGCTTTGGAATATCTAACAACTTCCACTCTCTGATGAGGACCAAC 1342
Qy      1321 AAGCTATGCGAGTTCTCCCTGTTCAGAAACTATGCGAGAGAGACTCCGAAACAGACTCT 1380
Db      1343 AAGCTATGCGAGTTCTCCCTGTTCAGAAACTATGCGAGAGAGACTCCGAAACAGACTCT 1402
Qy      1381 GATTAAGAGAGAACCCAGCATCCCAAGAGCTGACAGACCGCAGGCTTTCAGACAGCAG 1440
Db      1403 GATTAAGAGAGAACCCAGCATCCCAAGAGCTGACAGACCGCAGGCTTTCAGACAGCAG 1462
Qy      1441 AGCAAGCATTTGATTTCTGCTCAGAACAGAGAGAGTGGCACTCCGCAAGAGTCCCTTTTA 1500
Db      1463 AGCAAGCATTTGATTTCTGCTCAGAACAGAGAGAGTGGCACTCCGCAAGAGTCCCTTTTA 1522
Qy      1501 TCTCAGTGCATGGAAGTGGGAGCGTGGAGGACAAATTAACACACAGCTTCTTTTCGGC 1560
Db      1523 TCTCAGTGCATGGAAGTGGGAGCGTGGAGGACAAATTAACACACAGCTTCTTTTCGGC 1582
Qy      1561 CTTTTCACAGCCAGCAGACACTCAGAAATGCTGTGCGCTGGGCTTTAAGGCTGGCAC 1620
Db      1583 CTTTTCACAGCCAGCAGACACTCAGAAATGCTGTGCGCTGGGCTTTAAGGCTGGCAC 1642
Qy      1621 TCGGATATCTGGGCCCCCAGACCTTACCTTCTCTGACACAGAGCTGGTATTTTGGC 1680
Db      1643 TCGGATATCTGGGCCCCCAGACCTTACCTTCTCTGACACAGAGCTGGTATTTTGGC 1702
Qy      1681 ACAGAGTCCCTCAACTTTCTACTGCTGCTGACCATCTACAGAGGAGAGTGCAGTTACTCT 1740
Db      1703 ACAGAGTCCCTCAACTTTCTACTGCTGCTGACCATCTACAGAGGAGAGTGCAGTTACTCT 1762
Qy      1741 GCCTACAGCTGACCGCAGCTGCCACTGTGGAGAGCAAACTATTTCTGTGCGCAGGCGG 1800
Db      1763 GCCTACAGCTGACCGCAGCTGCCACTGTGGAGAGCAAACTATTTCTGTGCGCAGGCGG 1822
Qy      1801 CAGAAAGCAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATTAAGAGAGCCCTTTGAA 1860
Db      1823 CAGAAAGCAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATTAAGAGAGCCCTTTGAA 1882
Qy      1861 AAGCAGTTTAAAGCGAAGCTGCAAAATGGAATTTGAGAGAGCATCATGTCAAGAAC 1920
Db      1883 AAGCAGTTTAAAGCGAAGCTGCAAAATGGAATTTGAGAGAGCATCATGTCAAGAAC 1942
Qy      1921 AGGTCAAGGGAAGAGCTGGGGAAGAGTGGCAGTCACTTATTCGGGCAAGCATGGA 1980
Db      1943 AGGTCAAGGGAAGAGCTGGGGAAGAGTGGGAGTCACTTATTCGGGCAAGCATGGA 2002
Qy      1981 ATCATTGAGGTCTCC 1995
Db      2003 ATCATTGAGGTCTCC 2017

```

RESULT 13

```

US-10-072-012-257
; Sequence 257, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Speyrek, Kimberly
; APPLICANT: Zertusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Bisha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.

```

```
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Molenc, Adam R.
APPLICANT: Pena, Carol E. A.
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 257
LENGTH: 2071
TYPE: DNA
ORGANISM: Homo sapiens
US-10-072-012-257
```

Query Match 98.4%; Score 1962.2; DB 13; Length 2071;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

```
QY 1 ATGCCCATGAGATGATGGAATCTCAATTTTACTGAGAGTTGGCTTGTGGA 60
DB 61 ATGCCCATGAGATGATGGAATCTCAATTTTACTGAGAGTTGGCTTGTGGA 60
QY 61 AGTGAACGGAAGAAAGTGTCTTAATTGATGACCGGCATTGTGGAATCAATACATCC 120
DB 121 AGTGAACGGAAGAAAGTGTCTTAATTGATGACCGGCATTGTGGAATCAATACATCC 120
QY 121 CACATTTTGAAGCATTAATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 180
DB 181 CACATTTTGAAGCATTAATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 180
QY 181 GACAAAGTTTAATTAGAGACTCATCAGATTCAAGCAACATAAGTTGCACTTAT 240
DB 241 GACAAAGTTTAATTAGAGACTCATCAGATTCAAGCAACATAAGTTGCACTTAT 240
QY 241 TGCGATCGAAGGTTGATGTTACATCAAGCTCCCAAGATGTTCCCTCTCTTCA 300
DB 301 TGCGATCGAAGGTTGATGTTACATCAAGCTCCCAAGATGTTCCCTCTCTTCA 300
QY 301 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGCTTCACTGTTCACTG 360
DB 361 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGCTTCACTGTTCACTG 360
QY 361 CTTCGAGGTGGTTGCTGAGTTCTCTGTTTCCCTGGCTCTGGAAGAAATCC 420
DB 420 CTTCGAGGTGGTTGCTGAGTTCTCTGTTTCCCTGGCTCTGGAAGAAATCC 420
```

```
DB 421 CTTCGAGGTGGTTGCTGAGTTCTCTGTTTCCCTGGCTCTGGAAGAAATCC 480
QY 421 ACTCTAGTCCCTAAGCTGATTTCTGAGCTTGTCTTAACTGTGCAATTTGGCCAAAC 480
DB 481 ACTTAGTCCCTAAGCTGATTTCTGAGCTTGTCTTAACTGTGCAATTTGGCCAAAC 540
QY 481 CGAATTTCCCAATCTTATCTTGGCTGAGAGATGCTCTCAACAAGAGCTGATA 540
DB 541 CGAATTTCCCAATCTTATCTTGGCTGAGAGATGCTCTCAACAAGAGCTGATA 600
QY 541 CAGCAAGATGAGATGTTGTTATGTTAAATCCAGCTAATACCTGCAAGCTCACTT 600
DB 601 CAGCAAGATGAGATGTTGTTATGTTAAATCCAGCTAATACCTGCAAGCTCACTT 660
QY 601 ATCCCGAGTCAATTTCTGCGGTCCTGGAATGACAGCTTTTGTGAGAAATTTTG 660
DB 661 ATCCCGAGTCAATTTCTGCGGTCCTGGAATGACAGCTTTTGTGAGAAATTTTG 720
QY 661 CCGTGTGGAACAAATCAGATTTTCAATTGAGAAAGCAAGCTCCCAATGATGTT 720
DB 721 CCGTGTGGAACAAATCAGATTTTCAATTGAGAAAGCAAGCTCCCAATGATGTT 780
QY 721 CTAGTCACTGTTTAACTGGAGATCTCCGCTCCGCAATGCTTATGCTTATCATG 780
DB 781 CTAGTCACTGTTTAACTGGAGATCTCCGCTCCGCAATGCTTATGCTTATCATG 840
QY 781 AAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAGAACTT 840
DB 841 AAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAGAACTT 900
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGGAATGAGAAAGAAAGAACTT 900
DB 901 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGGAATGAGAAAGAAAGAACTT 960
QY 901 CAGCTGAGACATGAGGCTCAAGAGCAAACTCAAGCTGCTGACCTGAGAGCAAT 960
DB 961 CAGCTGAGACATGAGGCTCAAGAGCAAACTCAAGCTGCTGACCTGAGAGCAAT 1020
QY 961 GAACTGTGCTGTCTGCTGATGAGAGGTTGAGCAAGAAAGCAAGCTCTGATCAAC 1020
DB 1021 GAACTGTGCTGTCTGCTGATGAGAGGTTGAGCAAGAAAGCAAGCTCTGATCAAC 1080
QY 1021 TGTGCGACCTGTGCTGCTGATGAGAGGTTGAGCAAGAAAGCAAGCTCTGATCAAC 1080
DB 1081 TGTGCGACCTGTGCTGCTGATGAGAGGTTGAGCAAGAAAGCAAGCTCTGATCAAC 1132
QY 1081 CCCAGCTGCTGCTGCTGATGAGAGGTTGAGCAAGAAAGCAAGCTCTGATCAAC 1140
DB 1133 CCCAGCTGCTGCTGCTGATGAGAGGTTGAGCAAGAAAGCAAGCTCTGATCAAC 1191
QY 1133 AGTGGCTGCTGCTGCTGATGAGAGGTTGAGCAAGAAAGCAAGCTCTGATCAAC 1200
DB 1192 AGTGGCTGCTGCTGCTGATGAGAGGTTGAGCAAGAAAGCAAGCTCTGATCAAC 1251
QY 1201 TGTGCTGATTAATCAATGATTTCAATTTAGAGCAAGCAATGATCAAGCTCTT 1260
DB 1252 TGTGCTGATTAATCAATGATTTCAATTTAGAGCAAGCAATGATCAAGCTCTT 1311
QY 1261 TGTGCTGATTAATCAATGATTTCAATTTAGAGCAAGCAATGATCAAGCTCTT 1320
DB 1312 TGTGCTGATTAATCAATGATTTCAATTTAGAGCAAGCAATGATCAAGCTCTT 1371
QY 1321 TGTGCTGATTAATCAATGATTTCAATTTAGAGCAAGCAATGATCAAGCTCTT 1380
DB 1372 TGTGCTGATTAATCAATGATTTCAATTTAGAGCAAGCAATGATCAAGCTCTT 1431
QY 1381 TGTGCTGATTAATCAATGATTTCAATTTAGAGCAAGCAATGATCAAGCTCTT 1440
DB 1432 TGTGCTGATTAATCAATGATTTCAATTTAGAGCAAGCAATGATCAAGCTCTT 1491
QY 1441 TGTGCTGATTAATCAATGATTTCAATTTAGAGCAAGCAATGATCAAGCTCTT 1500
DB 1492 TGTGCTGATTAATCAATGATTTCAATTTAGAGCAAGCAATGATCAAGCTCTT 1551
```

Qy	1501	TCTCCACATCGAAGTGGAGCGTGGAGACAATTACACACAGCTTCTTTTGGC	1560
Db	1552	TCTCCACATCGAAGTGGAGCGTGGAGACAATTACACACAGCTTCTTTTGGC	1611
Qy	1561	CTTTCACACGACGACGACCTCTACGAAGTCTGTGGCTTGCGCTTTAAGGGCTGGCAC	1620
Db	1612	CTTTCACACACGACGACGACCTCTACGAAGTCTGTGGCTTGCGCTTTAAGGGCTGGCAC	1671
Qy	1621	TCGGATATCTTGGGCCCCCAGACCTCTACCCCTTCCCGACACAGAGCTGGTATTTTGGC	1680
Db	1672	TCGGATATCTTGGGCCCCCAGACCTCTACCCCTTCCCGACACAGAGCTGGTATTTTGGC	1731
Qy	1681	ACAGAGTCTTCACACTTCTACTCTTGCGCTACGCCATCTACGAGGCGAGTGGCAATTACTCT	1740
Db	1732	ACAGAGTCTTCACACTTCTACTCTTGCGCTACGCCATCTACGAGGCGAGTGGCAATTACTCT	1791
Qy	1741	GCCTACAGCTGCAGCGACGTGCCACTTGGGGAGACCAAGTCTATTCTGTGCGGAGGCGG	1800
Db	1792	GCCTACAGCTGCAGCGACGTGCCACTTGGGGAGACCAAGTCTATTCTGTGCGGAGGCGG	1851
Qy	1801	CAGAAGCCAAGTGAACAGAGCTGACTCGCGGGCGAGCTGGCATGAAGAGAGCCCTTTGAA	1860
Db	1852	CAGAAGCCAAGTGAACAGAGCTGACTCGCGGGCGAGCTGGCATGAAGAGAGCCCTTTGAA	1911
Qy	1861	AAGGAGTTTAACCGAGAGGCTGCCAATTGGAAATTTGGAGAGACATCATATGACAGAAC	1920
Db	1912	AAGGAGTTTAACCGAGAGGCTGCCAATTGGAAATTTGGAGAGACATCATATGACAGAAC	1971
Qy	1921	AGGTCAACGGGAAAGCTGGGGAAAGTGGGCAGTCAAGTCTATTCTGGGCGAGCATGGAA	1980
Db	1972	AGGTCAACGGGAAAGCTGGGGAAAGTGGGCAGTCAAGTCTATTCTGGGCGAGCATGGAA	2031
Qy	1981	ATCATTAGAGTCTCC	1995
Db	2032	ATCATTAGAGTCTCC	2046

RESULT 14
US-10-072-012-255
Sequence 255, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernhuse, Bryan
APPLICANT: Patturajan, Meera
APPLICANT: Shimkets, Richard
LI, LI
APPLICANT: Gangolli, Esha
APPLICANT: Padigaru, Muraidhara
APPLICANT: Anderson, David W.
APPLICANT: Raastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsebrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: 01/10/072,012
PRIOR APPLICATION NUMBER: 60/265,102
CURRENT FILING DATE: 2002-01-30
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514

	Query Match	95.6%; Score 1906.8; DB 13; Length 2200;	
	Best Local Similarity	97.7%; Pred. No. 0;	
	Matches 1993; Conservative 0; Mismatches	2; Indels 45; Gaps	4
QY	1 ATGCCCCATGAGATGATTGGAACTCAATTTGTTACTGAGAGGTGGTGGCTCTGCTGGAA	60	
DB	61 ATGCCCATGAGATGATTGGAACTCAATTTGTTACTGAGAGGTGGTGGCTCTGCTGGAA	120	
QY	61 AGTGAACGGAAAAAGTCTGCTAATTGATAGCCGGCCATTGTGTGGAATACATCATCC	120	
DB	121 AGTGAACGGAAAAAGTCTGCTAATTGATAGCCGGCCATTGTGTGGAATACATCATCC	180	
QY	121 CACATTTTGGAGGCACTTAATATCACTGCTCCAAAGCTTATGAAACGAAGTTGGCAAG	180	
DB	181 CACATTTTGGAGGCACTTAATATCACTGCTCCAAAGCTTATGAAACGAAGTTGGCAAG	240	
QY	181 GACAAAGTGTAAATTACAGAGCTATCCAGATTACGGAAACATA-----AG	228	
DB	241 GACAAAGTGTAAATTACAGAGCTATCCAGATTACGGAAACATAAGTAAACGCTCAG	300	
QY	229 GTTGAACATTTGATTCAGATCAGAGGTTGATTTACATCAAGATCCCAAGATTTGCC	288	
DB	301 GTTGAACATTTGATTCAGATCAGAGGTTGATTTACATCAAGATCCCAAGATTTGCC	360	
QY	289 TCTCTCTCTTCAGACTGTTTTCTCACTGTACTTCTGGGTAAACTGAGAAAGCTTCAAC	348	
DB	361 TCTCTCTCTTCAGACTGTTTTCTCACTGTACTTCTGGGTAAACTGAGAAAGCTTCAAC	420	
QY	349 TCTGTTCACTGCTTGC-----AGTGGGTTTCTGAGTTCTCTGTTGTTTC	396	
DB	421 TCTGTTCACTGCTTGCAGGTTTATCTGATGGTGGTTTCTGAGTTCTCTGTTGTTTC	480	
QY	397 CTTGGGCTCTGTGAAGAAAAATCCACTGTAAGTCCCTACCTGCAATTTCTCAACCTTGCTTA	456	
DB	481 CTTGGGCTCTGTGAAGAAAAATCCACTGTAAGTCCCTACCTGCAATTTCTCAACCTTGCTTA	540	
QY	457 CCTGTTCGCAACATTTGGGCGCAACCCGAATTTTCCCAATCTTTATCTTGGCTGCACGGA	516	
DB	541 CCTGTTCGCAACATTTGGGCGCAACCCGAATTTTCCCAATCTTTATCTTGGCTGCACGGA	600	
QY	517 GATGTCTCTCAACAGAGCTGATACAGACAGATGGATTTGTTATGTGTTAAATGCCAGC	576	
DB	601 GATGTCTCTCAACAGAGCTGATACAGACAGATGGATTTGTTATGTGTTAAATGCCAGC	660	
QY	577 TATATCTGTCCAAAGCCTGACTTATATCCCGAGTCTATTTCCGCGGTGGCCCTGTGAT	636	
DB	661 TATATCTGTCCAAAGCCTGACTTATATCCCGAGTCTATTTCCGCGGTGGCCCTGTGAT	720	

```

QY 637 GACAGCTTTTGTGAGAAATTTTGCCTGGTGTGACAAATCAGTAATTCAT----- 690
Db 721 GACAGCTTTTGTGAGAAATTTTGCCTGGTGTGACAAATCAGTAATTCATTTGGTAAG 780
QY 691 -----GAGAAAGCAAAAGCCTCCATATGATGTGTCTAGTGTGACGTGTAGT 738
Db 781 TTGACTTATACAGAGAAAGCAAAAGCCTCCATATGATGTGTCTAGTGTGACGTGTAGT 840
QY 739 GAGATCTCCGCTCCGACCATCGCTATGCTTACATCATGAGAGATGACATGTCT 798
Db 841 GAGATCTCCGCTCCGACCATCGCTATGCTTACATCATGAGAGATGACATGTCT 900
QY 799 TTATATGAGCTTAC---AGATTGTGAAAGAAAGAAAGACCTACTATATCTCCAAATCTC 855
Db 901 TTATATGAGCTTACAGAGAAATTTGTGAAAGAAAGAAAGACCTACTATATCTCCAAATCTC 960
QY 856 AATTTTCTGGCCAACTCTGAGACTATGAGAAAGAAATTTAAGAACAGACTGAGACATCA 915
Db 961 AATTTTCTGGCCAACTCTGAGACTATGAGAAAGAAATTTAAGAACAGACTGAGACATCA 1020
QY 916 GGGCCAAAGAGCAAACTCAAGCTGTGCACTGTGAGAAAGCAAAATGAACCTGTCCCTGCT 975
Db 1021 GGGCCAAAGAGCAAACTCAAGCTGTGCACTGTGAGAAAGCAAAATGAACCTGTCCCTGCT 1080
QY 976 GTCTCAGAGGAGTGTGACAGAAAGAGAGAGAGCCCTCACTGCTCACTGTGCTGCT 1035
Db 1081 GTCTCAGAGGAGTGTGACAGAAAGAGAGAGAGCCCTCACTGCTCACTGTGCTGCT 1140
QY 1036 ACCTCAGAGGAGTGTGACAGAAAGAGAGAGAGCCCTCACTGCTCACTGTGCTGCT 1095
Db 1141 ACCTCAGAGGAGTGTGACAGAAAGAGAGAGAGCCCTCACTGCTCACTGTGCTGCT 1200
QY 1096 GTGAGAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1155
Db 1201 GTGAGAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1156 TCCGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215
Db 1261 TCCGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1216 TCAGTTTCATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1275
Db 1321 TCAGTTTCATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1276 GCTTTGGAATATCAAAACCTTCACTACTGTGATGAGAGAGAGAGAGAGAGAGAGAG 1335
Db 1381 GCTTTGGAATATCAAAACCTTCACTACTGTGATGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1336 TCCCTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1395
Db 1441 TCCCTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1396 AGCATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1455
Db 1501 AGCATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1456 TCGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1515
Db 1561 TCGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1516 AGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1575
Db 1621 AGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1576 CAGACCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1635
Db 1681 CAGACCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 1636 CCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1695
Db 1741 CCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800

```

```

QY 1696 TTCTACTCTGCTCAAGCCATCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1755
Db 1801 TTCTACTCTGCTCAAGCCATCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
QY 1756 CAGTCCCACTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1815
Db 1861 CAGTCCCACTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 1816 AGAGCTGACTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1875
Db 1921 AGAGCTGACTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
QY 1876 AGAGCTGCAAAATGGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1935
Db 1981 AGAGCTGCCAAATGGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
QY 1936 CTGGGAAAGTGGGAGTCAAGTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1995
Db 2041 CTGGGAAAGTGGGAGTCAAGTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100

RESULT 15
US-09-964-277-20
; Sequence 20, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucite, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OR INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125_434
; CURRENT APPLICATION NUMBER: US/09/964,277
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PasterSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-20

Query Match      82.9%; Score 1653.8; DB 9; Length 3332;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1829; Conservative 0; Mismatches 2; Indels 164; Gaps 1;

QY 1 ATGGCCATGAGATGTGGAACCAATTTGTATCTGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 562 ATGGCCATGAGATGTGGAACCAATTTGTATCTGAGAGAGAGAGAGAGAGAGAGAGAG 621
QY 61 AGTGAACCGAAAAAGTGTCTAATTTAGTAGCCGCAATTTGTGAATACAAATATCC 120
Db 622 AGTGAACCGAAAAAGTGTCTAATTTAGTAGCCGCAATTTGTGAATACAAATATCC 120
QY 121 CACATTTGGAAGCCATTAATATCAAGCTCCAGCTTTATGAGCGGAAGTTGCAACAG 180
Db 682 CACATTTGGAAGCCATTAATATCAAGCTCCAGCTTTATGAGCGGAAGTTGCAACAG 180
QY 181 GACAAAGTGTATTTACAGAGCTCATTCAGATTCAGGAAACATTAAGTTGACATTTAT 240
Db 742 GACAAAGTGTATTTACAGAGCTCATTCAGATTCAGGAAACATTAAGTTGACATTTAT 240
QY 742 GACAAAGTGTATTTACAGAGCTCATTCAGATTCAGGAAACATTAAGTTGACATTTAT 240
Db 241 TGCAGTCAGAAAGTGTATTTAGATTAAGAGCTCCCAAGTGTGCTCTCTCTTCA 300
QY 802 TGCAGTCAGAAAGTGTATTTAGATTAAGAGCTCCCAAGTGTGCTCTCTCTTCA 360
Db 301 GACTGTTTTCTCACTGTATCTTGGGTAAACTGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 862 GACTGTTTTCTCACTGTATCTTGGGTAAACTGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 CTGAGAGTGGGTTTGTGAGATTTCTGCTGTTTCTGAGAGAGAGAGAGAGAGAGAGAG 420
QY 922 CTGAG----- 926
Db 421 ACTTATGCTTACTGAGATTTCTAGAGCTTGTCTACTGTTGCAACATTTGGGCAAC 480

```

Db 927 ----- 926
Qy 481 GGAATTCCTCCAACTTTATCTTGCTGCAGAGAGTCTCTCAAGAGAGCTGATA 540
Db 927 ----- AGGAGCTGATG 937
Qy 541 CAGCAGAAATGGGATGGTATATGTTTAAATGCCAGCTATACTGTCCAAAGCCTGACTTT 600
Db 938 CACACAGAAATGGGATGGTATATGTTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTT 997
Qy 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG 660
Db 998 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG 1057
Qy 661 CCGTGTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTTCATATGATGTT 720
Db 1058 CCGTGTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTTCATATGATGTT 1117
Qy 721 CTAGTGACCTGTTAGCTGGGATCTCCGCTCCGCACTATGCTATGCTATCATCATG 780
Db 1118 CTAGTGACCTGTTAGCTGGGATCTCCGCTCCGCACTATGCTATGCTATCATCATG 1177
Qy 781 AAGAGATGACATGTCTTAGATGAAGCTTACAGATTTGTGAAAAGAAAAGCACTACT 840
Db 1178 AAGAGATGACATGTCTTAGATGAAGCTTACAGATTTGTGAAAAGAAAAGCACTACT 1237
Qy 841 ATATCTCCAACTTCAATTTTCTTGCGCCAACTCTGTGACTATGAGAAAGATTAAGAAC 900
Db 1238 ATATCTCCAACTTCAATTTTCTTGCGCCAACTCTGTGACTATGAGAAAGATTAAGAAC 1297
Qy 901 CAGACTGGACATCAAGGCGCAAAAGAGCAAACTCAAGCTGTGACCTGTGAGAAAGCCAAAT 960
Db 1298 CAGACTGGACATCAAGGCGCAAAAGAGCAAACTCAAGCTGTGACCTGTGAGAAAGCCAAAT 1357
Qy 961 GAACCTGTCCCTGCTGTCTAGAGAGGTGACAGAAAAGCAGAGCCCTCACTGACCAACC 1020
Db 1358 GAACCTGTCCCTGCTGTCTAGAGAGGTGACAGAAAAGCAGAGCCCTCACTGACCAACC 1417
Qy 1021 TGTGCCGACTCTGTCTACCTTAGAGGACAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
Db 1418 TGTGCCGACTCTGTCTACCTTAGAGGACAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1477
Qy 1081 CCCAGCGTGCAGCGTGCAGCGCTGCTGTGAGAGACAGCCCGCTGTGATCAGGCGCTC 1140
Db 1478 CCCAGCGTGCAGCGTGCAGCGCTGCTGTGAGAGACAGCCCGCTGTGATCAGGCGCTC 1537
Qy 1141 AGTGGCTGACCTGTCCGACAGACAGGCTGAAAGACAGCAATTAAGCTCAAGCGTTCCTTC 1200
Db 1538 AGTGGCTGACCTGTCCGACAGACAGGCTGAAAGACAGCAATTAAGCTCAAGCGTTCCTTC 1597
Qy 1201 TCTCTGGATATCAAAATCAGTTTCAATTTACGCCAGCATGGCAGCATCTTACATGGCTTC 1260
Db 1598 TCTCTGGATATCAAAATCAGTTTCAATTTACGCCAGCATGGCAGCATCTTACATGGCTTC 1657
Qy 1261 TCTCTCATCAAAAGATGCTTTGGAATATACTCAAACTTCACTACTGTGATGGGACCAAC 1320
Db 1658 TCTCTCATCAAAAGATGCTTTGGAATATACTCAAACTTCACTACTGTGATGGGACCAAC 1717
Qy 1321 AAGCTATGCGAGTTCTCCCGTGTTCAGAACTATCGAGAGAGATCCCGAAACCAAGTCTT 1380
Db 1718 AAGCTATGCGAGTTCTCCCGTGTTCAGAACTATCGAGAGAGATCCCGAAACCAAGTCTT 1777
Qy 1381 GATTAAGAGAGAGCAGACATCCCGAAGAGCTGACAGCCGACAGGCTTTCAGACAGCCAG 1440
Db 1778 GATTAAGAGAGAGCAGACATCCCGAAGAGCTGACAGCCGACAGGCTTTCAGACAGCCAG 1837
Qy 1441 AGCAAGGATTTGCAATTTGGTCAAGAACAGACAGAGTGGCAACCGCCAGAGAGTCCCTTTTA 1500
Db 1838 AGCAAGGATTTGCAATTTGGTCAAGAACAGACAGAGTGGCAACCGCCAGAGAGTCCCTTTTA 1897
Qy 1501 TCTCACTGATCGAAGTGGAGAGCTGAGAGACAAATTAACACACAGCTTCTTTTCGCGC 1560

Db 1898 TCTCACTGATCGAAGTGGAGAGCTGAGAGACAAATTAACACACAGCTTCTTTTCGCGC 1957
Qy 1561 CTTTCCACAGAGCAGAGACCTCAAGAGTCTGTGGGCTTGAGGCTTGAGGCTTGAGC 1620
Db 1958 CTTTCCACAGAGCAGAGACCTCAAGAGTCTGTGGGCTTGAGGCTTGAGGCTTGAGC 2017
Qy 1621 TCGATATCTTGAGGCCCCCAGAGCTTACCCCTCCCTGACACAGAGCTGTATTTTGGC 1680
Db 2018 TCGATATCTTGAGGCCCCCAGAGCTTACCCCTCCCTGACACAGAGCTGTATTTTGGC 2077
Qy 1681 ACAGAGTCTCACTTCTACTCTGCTCAGCCATCTACGAGGACAGTCCAGTTACTCT 1740
Db 2078 ACAGAGTCTCACTTCTACTCTGCTCAGCCATCTACGAGGACAGTCCAGTTACTCT 2137
Qy 1741 GCCTTAAGCTGACAGCTGCCCCATTTGCGAGAACAGTCTATTTCTGTGCGCAGGCGG 1800
Db 2138 GCCTTAAGCTGACAGCTGCCCCATTTGCGAGAACAGTCTATTTCTGTGCGCAGGCGG 2197
Qy 1801 CAGAAAGCAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATGAAAGAGAGCCCTTTGAA 1860
Db 2198 CAGAAAGCAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATGAAAGAGAGCCCTTTGAA 2257
Qy 1861 AAGCAGTTTAAAGCAGAAAGCTGCCAAATGAAATTTGAGAGAGCATGTGAGAAAC 1920
Db 2258 AAGCAGTTTAAAGCAGAAAGCTGCCAAATGAAATTTGAGAGAGCATGTGAGAAAC 2317
Qy 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTAGCTTTTGGGAGAGATGAA 1980
Db 2318 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTAGCTTTTGGGAGAGATGAA 2377
Qy 1981 ATCATTAGAGTCTCC 1995
Db 2378 ATCATTAGAGTCTCC 2392

Search completed: June 22, 2004, 03:25:18
Job time : 1198.09 secs

~~... This Page Blank (aspj0)~~

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 17:08:00 ; Search time 165.047 Seconds
(Without alignments)
6707.969 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccatgagatgatggtg.....tggaatcatgtgagttctcc 1995

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCUTS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1991.8	99.8	1998 4	US-09-816-494-3 Sequence 3, Appl1
2	1991.8	99.8	3544 4	US-09-816-494-1 Sequence 1, Appl1
3	387.8	19.4	2377 4	US-09-820-668-3 Sequence 91, Appl1
4	223	11.2	279 4	US-09-016-434-91 Sequence 1, Appl1
5	124.6	6.2	1830 4	US-09-557-921-1 Sequence 3, Appl1
6	106.8	5.4	2303 4	US-09-922-146-3 Sequence 1135, Ap
7	96.2	4.8	2109 4	US-09-016-434-1135 Sequence 946, App
8	96.2	4.7	1208 4	US-09-023-655-347 Sequence 801, App
9	94.6	4.7	1619 4	US-09-702-705-801 Sequence 801, App
10	93.6	4.7	1619 4	US-09-736-457-801 Sequence 801, App
11	93.6	4.7	1619 4	US-09-614-1248-801 Sequence 801, App
12	93.6	4.7	1619 4	US-09-614-1248-801 Sequence 801, App
13	93.6	4.7	1619 4	US-09-614-1248-801 Sequence 801, App
14	93.6	4.7	1619 4	US-09-589-184-801 Sequence 801, App
15	93.6	4.7	1619 4	US-09-589-184-801 Sequence 801, App
16	93.6	4.7	1619 4	US-09-589-184-801 Sequence 801, App
17	93.6	4.7	1619 4	US-09-589-184-801 Sequence 801, App
18	93.6	4.7	1619 4	US-09-589-184-801 Sequence 801, App
19	93.6	4.7	1619 4	US-09-589-184-801 Sequence 801, App
20	93.6	4.7	1619 4	US-09-589-184-801 Sequence 801, App
21	92	4.6	1238 4	US-08-530-290-11 Sequence 11, Appl
22	92	4.6	1238 4	US-09-702-705-803 Sequence 803, App
23	92	4.6	1238 4	US-09-736-457-803 Sequence 803, App
24	92	4.6	1238 4	US-09-614-1248-803 Sequence 803, App
25	92	4.6	1238 4	US-09-614-1248-803 Sequence 803, App
26	92	4.6	1238 4	US-09-614-1248-803 Sequence 803, App
27	92	4.6	1238 4	US-09-614-1248-803 Sequence 803, App

28	92	4.6	2064 4	US-09-614-1248-825 Sequence 825, App
29	92	4.6	2064 4	US-09-671-325-825 Sequence 825, App
30	92	4.6	2064 4	US-09-589-184-825 Sequence 825, App
31	92	4.6	2109 4	US-09-702-705-825 Sequence 826, App
32	92	4.6	2109 4	US-09-736-457-826 Sequence 826, App
33	92	4.6	2109 4	US-09-614-1248-826 Sequence 826, App
34	92	4.6	2109 4	US-09-671-325-826 Sequence 826, App
35	92	4.6	2109 4	US-09-589-184-826 Sequence 826, App
36	92	4.6	2240 4	US-09-016-434-1100 Sequence 1100, App
37	89.4	4.5	2000 4	US-09-016-434-1281 Sequence 1291, App
38	86.2	4.3	240 4	US-09-016-434-776 Sequence 776, App
39	85.6	4.3	1987 2	US-08-990-379-1 Sequence 1, Appl1
40	85.6	4.3	1993 2	US-08-990-379-2 Sequence 2, Appl1
41	80.4	4.0	944 4	US-09-371-671B-10 Sequence 10, Appl
42	79	4.0	539 4	US-09-389-681-311 Sequence 311, App
43	79	4.0	539 4	US-09-620-405B-311 Sequence 311, App
44	79	4.0	539 4	US-09-339-338-311 Sequence 311, App
45	79	4.0	539 4	US-09-433-826B-311 Sequence 311, App

ALIGNMENTS

RESULT 1				
US-09-816-494-3				
Sequence 3, Application US/09816494				
Patent No. 6664089				
GENERAL INFORMATION:				
APPLICANT: Meyers, Rachel A.				
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY				
TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR				
FILE REFERENCE: 10448-030002				
CURRENT APPLICATION NUMBER: US/09/816,494				
CURRENT FILING DATE: 2001-03-23				
PRIOR APPLICATION NUMBER: US 60/191,858				
PRIOR FILING DATE: 2000-03-24				
NUMBER OF SEQ ID NOS: 10				
SOFTWARE: FastSeq for Windows Version 4.0				
SEQ ID NO 3				
LENGTH: 1998				
TYPE: DNA				
ORGANISM: Homo sapiens				
US-09-816-494-3				
Query Match				
Best Local Similarity 99.8%; Score 1991.8; DB 4; Length 1998;				
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
QY	1	ATGGCCCATGAGATGATGGAATCTCAATTTGTAAGAGGTTGGCTGCTGGAA	60	
DB	1	ATGGCCCATGAGATGATGGAATCTCAATTTGTAAGAGGTTGGCTGCTGGAA	60	
QY	61	AGTGAACGGAAGAGTGTCTGATTTGATGATGAGCGGCAATTTGGAATCAATATCC	120	
DB	61	AGTGAACGGAAGAGTGTCTGATTTGATGATGAGCGGCAATTTGGAATCAATATCC	120	
QY	121	CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATGAAGAGTTGCAACAG	180	
DB	121	CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATGAAGAGTTGCAACAG	180	
QY	181	GACAAAGTGTATTCAGAGCTCATCCAGCATTCAGCGAATCAATAGTTGACATTGAT	240	
DB	181	GACAAAGTGTATTCAGAGCTCATCCAGCATTCAGCGAATCAATAGTTGACATTGAT	240	
QY	241	TGCAATCAGAGGTTGATTTACGATCAAAAGCTCCCAAGATGCTCTCTCTTCA	300	
DB	241	TGCAATCAGAGGTTGATTTACGATCAAAAGCTCCCAAGATGCTCTCTCTTCA	300	
QY	301	GACTGTTTCTCCTGTAATCTTGGGTAATCTGGAAGAGCTTCAACTGTTCACCTG	360	
DB	301	GACTGTTTCTCCTGTAATCTTGGGTAATCTGGAAGAGCTTCAACTGTTCACCTG	360	
QY	361	CTTGCAAGTGGTGTCTGATGATCTCTGTTTCTCTGCTCTGTAAGGAATATCC	420	

Db 361 CTTGAGAGTGGTTTGGAGTGTCTCGTGTTCCTCGGCTCTGTGAGAGAAATCC 420
Qy 421 ACTCTAGTCCCTTACCTGATTTCTCAGCCTTGTACCTGTTCGCAACATTTGGGCCAAC 480
Db 421 ACTCTAGTCCCTTACCTGATTTCTCAGCCTTGTACCTGTTCGCAACATTTGGGCCAAC 480
Qy 481 CGAATTTCTCCCATCTTTATCTTGGTGTGCGCAGCGAGATGTCTCAACAGAGAGCTGATA 540
Db 481 CGAATTTCTCCCATCTTTATCTTGGTGTGCGCAGCGAGATGTCTCAACAGAGAGCTGATA 540
Qy 541 CAGAGAAATGGATGGTATGTATGTAAATGCGAGCTATACCTGTCCAAAGCTGACCTT 600
Db 541 CAGAGAAATGGATGGTATGTATGTAAATGCGAGCTATACCTGTCCAAAGCTGACCTT 600
Qy 601 ATCCCGAGTCTCATTTCCGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db 601 ATCCCGAGTCTCATTTCCGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Qy 661 CCGTGTGTGACAAATCAGTAGATTTCATTGAGAAAGAAAAGCTCCAAATGAGTGT 720
Db 661 CCGTGTGTGACAAATCAGTAGATTTCATTGAGAAAGAAAAGCTCCAAATGAGTGT 720
Qy 721 CTAGTGAAGTGTGTGCTGGGATCTCCGCTCCGCAACATGCTATGCTGATCATCATG 780
Db 721 CTAGTGAAGTGTGTGCTGGGATCTCCGCTCCGCAACATGCTATGCTGATCATCATG 780
Qy 781 AAGAGATGACATGTCTTTATGATGAAGCTTACAGATTTGTGAAAGAAAAGACCTACT 840
Db 781 AAGAGATGACATGTCTTTATGATGAAGCTTACAGATTTGTGAAAGAAAAGACCTACT 840
Qy 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGGAATATGAGAAAGATTAAAGAC 900
Db 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGGAATATGAGAAAGATTAAAGAC 900
Qy 901 CAGACTGAGATGAGGGCCAAAGAGCAAACTCAAGCTGCTCACTGTGAGAGAGCAAT 960
Db 901 CAGACTGAGATGAGGGCCAAAGAGCAAACTCAAGCTGCTCACTGTGAGAGAGCAAT 960
Qy 961 GAACCTGTCCCTGTCTCTCAGAGGGTGAAGAAAAGGAGAGCCCTCGATGCAACC 1020
Db 961 GAACCTGTCCCTGTCTCTCAGAGGGTGAAGAAAAGGAGAGCCCTCGATGCAACC 1020
Qy 1021 TGTGCGGACTGTCTCTCTCTCAGAGGGTGAAGAAAAGGAGAGCCCTCGATGCAACC 1080
Db 1021 TGTGCGGACTGTCTCTCTCTCAGAGGGTGAAGAAAAGGAGAGCCCTCGATGCAACC 1080
Qy 1081 CCCAGGTCGCCAGGTCGAGCGTCGCTGTGAAGAGCAAGCCGCTGTGAAGAGCGCTC 1140
Db 1081 CCCAGGTCGCCAGGTCGAGCGTCGCTGTGAAGAGCAAGCCGCTGTGAAGAGCGCTC 1140
Qy 1141 AGTGGGCTGCACTCTCCGAGAGAGAGCTGGAAGAGCAATAGCTCAAGCGTTCCTTC 1200
Db 1141 AGTGGGCTGCACTCTCCGAGAGAGAGCTGGAAGAGCAATAGCTCAAGCGTTCCTTC 1200
Qy 1201 TCTCTGATATCAATCAGTTTCAATTTCAAGCGAGATGCGAGATCTTAACTGCTTC 1260
Db 1201 TCTCTGATATCAATCAGTTTCAATTTCAAGCGAGATGCGAGATCTTAACTGCTTC 1260
Qy 1261 TCTCTGATATCAATCAGTTTCAATTTCAAGCGAGATGCGAGATCTTAACTGCTTC 1320
Db 1261 TCTCTGATATCAATCAGTTTCAATTTCAAGCGAGATGCGAGATCTTAACTGCTTC 1320
Qy 1321 AAGCTATGCAATGCTTCTGTTGAGAACTATGCAAACTTTCACATCTGTGATGGAGCAAC 1380
Db 1321 AAGCTATGCAATGCTTCTGTTGAGAACTATGCAAACTTTCACATCTGTGATGGAGCAAC 1380
Qy 1381 GATTAAG 1440
Db 1381 GATTAAG 1440
Qy 1441 AGGAG 1500
Db 1441 AGGAG 1500

Db 1441 AGGAG 1500
Qy 1501 TCTCTGATATGCAATGCTTCTGTTGAGAACTATGCAAACTTTCACATCTGTGATGGAGCAAC 1560
Db 1501 TCTCTGATATGCAATGCTTCTGTTGAGAACTATGCAAACTTTCACATCTGTGATGGAGCAAC 1560
Qy 1561 CTTTCCAG 1620
Db 1561 CTTTCCAG 1620
Qy 1621 TGGATATCTGGGCTCCGAG 1680
Db 1621 TGGATATCTGGGCTCCGAG 1680
Qy 1681 ACAGAGTCTCAGACTTCTACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1681 ACAGAGTCTCAGACTTCTACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Qy 1741 GCCTAAGCTGAG 1800
Db 1741 GCCTAAGCTGAG 1800
Qy 1801 CAG 1860
Db 1801 CAG 1860
Qy 1861 AAGAGATTAAG 1920
Db 1861 AAGAGATTAAG 1920
Qy 1921 AGGTCAAG 1980
Db 1921 AGGTCAAG 1980
Qy 1981 ATCATTTAGAGTCTCC 1995
Db 1981 ATCATTTAGAGTCTCC 1995

RESULT 2
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)...(2583)
US-09-816-494-1

Query Match 99.8%; Score 1991.8; DB 4; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGCGCCATGAGATGATGGAATCAATTTGTTACTGAGAGGTTGGCTGTGAGAA 60
Db 589 ATGCGCCATGAGATGATGGAATCAATTTGTTACTGAGAGGTTGGCTGTGAGAA 60
Qy 61 AGTGAACGAAAAAGTGTCTTAATTGATAGCGGCGCTTTGTGAATCAATCATCC 120
Db 61 AGTGAACGAAAAAGTGTCTTAATTGATAGCGGCGCTTTGTGAATCAATCATCC 120

Dh 649 AGTGAACGAAAAAGTCTGCTAATTGATAGCCGGCCATTGTGGAATACATATCC 708
Qy 121 CACATTTTGGAGCCATTAAATATCACTGCTCCAGTTTATGAAAGGAGTTGCAACAG 180
Dh 709 CACATTTTGGAGCCATTAAATATCACTGCTCCAGTTTATGAAAGGAGTTGCAACAG 768
Qy 181 GACAAAGTGTAAATTCAGAGCTCATCAGCATTTGAGGAAACAATAGGTGACATTGAT 240
Dh 769 GACAAAGTGTAAATTCAGAGCTCATCAGCATTTGAGGAAACAATAGGTGACATTGAT 828
Qy 241 TGCAGTCAGAGGTGTAGTTTACGATCAAAAGCTCCAGAGATGTGCTCTCTTCA 300
Dh 829 TGCAGTCAGAGGTGTAGTTTACGATCAAAAGCTCCAGAGATGTGCTCTCTTCA 888
Qy 301 GACTGTTTTCTCACTGTACTTCTGGGTAAACTGGAGAAAGCTTCAACTCTGCTC 360
Dh 889 GACTGTTTTCTCACTGTACTTCTGGGTAAACTGGAGAAAGCTTCAACTCTGCTC 948
Qy 361 CTGGCAGGTGGGTTTGCTGAGTTCTCTGCTGTTTCCCTGGCCCTGTGAAAGGAAATCC 420
Dh 949 CTGGCAGGTGGGTTTGCTGAGTTCTCTGCTGTTTCCCTGGCCCTGTGAAAGGAAATCC 1008
Qy 421 ACTTAGTCCCTACCTGCACTTCTCAGCCCTTGCTTACCTGTGCAACATTTGGGCCAAC 480
Dh 1009 ACTTAGTCCCTACCTGCACTTCTCAGCCCTTGCTTACCTGTGCAACATTTGGGCCAAC 1068
Qy 481 CGAATTTCTCCAACTTTATCTTGGCTGCCAGGAGATGCTCTCAACAGAGAGCTGATA 540
Dh 1069 CGAATTTCTCCAACTTTATCTTGGCTGCCAGGAGATGCTCTCAACAGAGAGCTGATA 1128
Qy 541 CACACAGATGGGATGGTTATGTGTTAAATGCAAGCATATCCTGCCAAAGCCGTGACTTT 600
Dh 1129 CACACAGATGGGATGGTTATGTGTTAAATGCAAGCATATCCTGCCAAAGCCGTGACTTT 1188
Qy 601 ATCCCGAGTCTGATTTCTGCGTGTGCTGTGATGACAGCTTTGTGAAATAATTTTG 660
Dh 1189 ATCCCGAGTCTGATTTCTGCGTGTGCTGTGATGACAGCTTTGTGAAATAATTTTG 1248
Qy 661 CCGTGTGGACAAATCAGTAGATTTTCAATTGAGAAAGCAAAAGCCTTCAATGATGTT 720
Dh 1249 CCGTGTGGACAAATCAGTAGATTTTCAATTGAGAAAGCAAAAGCCTTCAATGATGTT 1308
Qy 721 CTGTGACACTTTAGTGGGATCTCCGCTCCGCCCACTGCTATCGCTACATGATG 780
Dh 1309 CTGTGACACTTTAGTGGGATCTCCGCTCCGCCCACTGCTATCGCTACATGATG 1368
Qy 781 AAGAGATGACATGCTTAGATGAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 840
Dh 1369 AAGAGATGACATGCTTAGATGAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 1428
Qy 841 ATATCTCAAACTTCAATTTTCTGGGCCAACTCTGTGACTATGAGAAAGATTAAGAAC 900
Dh 1429 ATATCTCAAACTTCAATTTTCTGGGCCAACTCTGTGACTATGAGAAAGATTAAGAAC 1488
Qy 901 CAGACTGAGACATCAAGGCCAAAGAGAACTCAAGCTGTGACCTGGAGAAAGCCAAAT 960
Dh 1489 CAGACTGAGACATCAAGGCCAAAGAGAACTCAAGCTGTGACCTGGAGAAAGCCAAAT 1548
Qy 961 GAACCTGTCTGCTGTCTAGAGAGGTGACAGAAAGAGAGAGCCCTCACTGACACC 1020
Dh 1549 GAACCTGTCTGCTGTCTAGAGAGGTGACAGAAAGAGAGAGCCCTCACTGACACC 1080
Qy 1021 TGTGCGCACTCTGTACTCAGAGAGCAGAGCAAAAGGCCCTGTGATCCGCCAGGCTG 1080
Dh 1609 TGTGCGCACTCTGTACTCAGAGAGCAGAGCAAAAGGCCCTGTGATCCGCCAGGCTG 1668
Qy 1081 CCGAGCGTGGCCAGCGGTGACAGCCGTGTTAGAGACAGCCCGCTGTGATCAGCGCTC 1140
Dh 1669 CCGAGCGTGGCCAGCGGTGACAGCCGTGTTAGAGACAGCCCGCTGTGATCAGCGCTC 1728
Qy 1141 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCTTCTTTC 1200
Dh 1729 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCTTCTTTC 1788

Qy 1201 TCTCTGGATATCAAAATCAGTTTCAATATTCAGCCAGCATGGCAGCATCTTACATGGCTTC 1260
Dh 1789 TCTCTGGATATCAAAATCAGTTTCAATATTCAGCCAGCATGGCAGCATCTTACATGGCTTC 1848
Qy 1261 TCTCTCATCAGAAAGATGCTTTGGAAATCTACAAACCTTCCACTACCTGTGATGGGACCAAC 1320
Dh 1849 TCTCTCATCAGAAAGATGCTTTGGAAATCTACAAACCTTCCACTACCTGTGATGGGACCAAC 1308
Qy 1321 AAGCTATGCCAGTTCTCCCTGTTCAGAACTATGAGAGACACTCCGAAACCAATGCTT 1380
Dh 1909 AAGCTATGCCAGTTCTCCCTGTTCAGAACTATGAGAGACACTCCGAAACCAATGCTT 1368
Qy 1381 GATTAAGAGAAAGCCAGCATTCGCCAAGAGCTGCAACCGCCAGGCTTTCAGACAGCAG 1440
Dh 1969 GATTAAGAGAAAGCCAGCATTCGCCAAGAGCTGCAACCGCCAGGCTTTCAGACAGCAG 2028
Qy 1441 AGCAAGCAGATTGCAATTCGGTCAAGAACAGCAGCAGTGGCACCGCCAGAGGTCCCTTTTA 1500
Dh 2029 AGCAAGCAGATTGCAATTCGGTCAAGAACAGCAGCAGTGGCACCGCCAGAGGTCCCTTTTA 2088
Qy 1501 TCTTCATCTGATTCGAAGTGGAGCGTGGAGAGACATTTACACACAGCTTCTTTTGGCC 1560
Dh 2089 TCTTCATCTGATTCGAAGTGGAGCGTGGAGAGACATTTACACACAGCTTCTTTTGGCC 2148
Qy 1561 CTTTCCACACAGCAGACAGACCTCAGAGATCTGCGCTGGGCTTTAAGGCTGGCAC 1620
Dh 2149 CTTTCCACACAGCAGACAGACCTCAGAGATCTGCGCTGGGCTTTAAGGCTGGCAC 2208
Qy 1621 TCGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGCC 1680
Dh 2209 TCGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGCC 2268
Qy 1681 ACAGAGTCTTCACTTCTAATCTGCTGCTGACCATTTAGGAGGAGAGTCCAGTTACTT 1740
Dh 2269 ACAGAGTCTTCACTTCTAATCTGCTGCTGACCATTTAGGAGGAGAGTCCAGTTACTT 2328
Qy 1741 GCCTAAGCTGACGACGAGCGCCCACTTGGGAGACCAAGTCTAATCTGCGGACAGCGG 1800
Dh 2329 GCCTAAGCTGACGACGAGCGCCCACTTGGGAGACCAAGTCTAATCTGCGGACAGCGG 2388
Qy 1801 CAGAACCCAAAGTACAGAGCTGACTCGCGCGAGGCTGGCATGAGAGAGCCCTTTGAA 1860
Dh 2389 CAGAACCCAAAGTACAGAGCTGACTCGCGCGAGGCTGGCATGAGAGAGCCCTTTGAA 2448
Qy 1861 AAGCAGTTTAAAGCAGAAAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCAAGAAC 1920
Dh 2449 AAGCAGTTTAAAGCAGAAAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCAAGAAC 2508
Qy 1921 AGGTCAAGGAGAAAGCTGGGAGAAAGTGGGCAATGCTTAACTTTTCTGGGCAAGATGAA 1980
Dh 2509 AGGTCAAGGAGAAAGCTGGGAGAAAGTGGGCAATGCTTAACTTTTCTGGGCAAGATGAA 2568
Qy 1981 ATCATTGAGGTCTCC 1995
Dh 2569 ATCATTGAGGTCTCC 2583

RESULT 3
US-09-920-668-3
; Sequence 3, Application US/09920668
; Patent No. 6482644
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; APPLICANT: Brett P. Monia
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 8 EXPRESSION
; FILE REFERENCE: RTS-0246
; CURRENT APPLICATION NUMBER: US/09/920,668
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 2377
; TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (135) ... (2012)
US-09-920-668-3

Query Match 19.4%; Score 387.8; DB 4; Length 2377;
Best Local Similarity 59.9%; Pred. No. 1.7e-119;
Matches 716; Conservative 0; Mismatches 452; Indels 27; Gaps 3;

```

QY 28 ATTGTTACTGAGAGGTTGGTGGCTTCTGTTGAAAAGTGAACGGAAGAAAGTGTCTAATT 87
DB 165 ATGATGACCAAGAACTGCGCAGCCTGCTGGGGGCGGCGCTGGGGGGCGCGTGGTCAATC 224
QY 88 GATACCGGCGCATTTTGTGAAATACAAATATCCCACTTTTGAAGCCATTAATATCAAC 147
DB 225 GACAGCCGCTCTCTGTTGAGTACACAGCTGGGCTGTGCTGAGCTCCGTCAACATCTGCG 284
QY 148 TGTCTCAAGCTTATGAAAGCGAGTTGCAACAGACAAAGTGTAAATACAGAGCTCATC 207
DB 285 TGTCTCAAGCTGTGAAAGCGGCGCTGCGACAGAGGCAAGTGAACATTGCGGAGCTATC 344
QY 208 CAGCATTCAGCGAAAATAAAGTTGACATTGATTGACATCAAGATTGTTACTTATGAT 267
DB 345 CAGCGGCTGACGACGAGCGAGTGAAGGCTACGAGACCAACAGAGTGTGTCTATGAC 404
QY 268 CAAGCTCCCAAGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 404
DB 405 CAGAGCAGCGGAGACCGCAGCGCTGCGCGCAACAGCTTCTCTCATCTCTGCTGAGC 464
QY 328 AAACGTGAAAGAGCTTCACTCTGTTCACTGCTGAGGTTGGTGTGAGTTCTCT 387
DB 465 AAGCTGAGCGCTGCTCTGACAGCTGCGCATCTCTGAGGCGCTTGGCACCTTCTCC 524
QY 388 CGTTGTTTCCCTGGGCTCTGTGAGAGAAATCCCT---CTAGTCCCTACCTGACTTCT 444
DB 525 TCTGCTTCCCGGCTCTGTGAGAGGCAAGCTGCTGCTCTGCTACCTGAGGCTCTCC 584
QY 445 CAGCGTCTTACCTGTTGGCAATGTTGGGCAACCGAATTTCTCCCAATCTTATCTT 504
DB 585 CAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644
QY 505 GCGTCCAGCAGATGTTCTCAACAGAGAGCTATACAGCAAGTGGATGGTTATGTG 564
DB 645 GGTCTGAGAGAGCGTCTTAACAGAGATGTATGACGCAAAAGATGAAGCTAGCTC 704
QY 565 TTAATGCAAGCTATACCTGTCAAGAGCTTATCCCGAGTCTTATTCCTGCT 624
DB 705 CTGAGCGCAAGCACTCTGCTCCCAAGCTGACTTCACTGTGAGAGCGCTTATGCTG 764
QY 625 GTGCTGTGATGACAGCTTTTGTGAAATTTTGGCGTGTGGAACAATCAGTATGAT 684
DB 765 GTCCCATCAAGCACTACTGTGAAAATCTGCTGCTGCTGCTGCAAGTCAATGAG 824
QY 685 TTCAATGAGAAAGAAAGCCCTCCATGATGTGTTAGTGCAGCTTTAGGCTGGGATC 744
DB 825 TTCAATGAGAAAGCCCAAGCTCTCCAGCTGCCAAGTCACTGCTGCTGCTGCTGCTG 884
QY 745 TCCCGTCCGCAACATGCTATGCGCTCAATCAATGAGAGATGACATGTTTGAAT 804
DB 885 TCCCGTCTGCAACATGCTGCTCAATCAATGAGAGCAATGCTGCTGCTGCTGCTGCTG 864
QY 805 GAAAGCTTACAGATTGTGAAAGAAAGAAAGCTTATATCTTCAACTTCAATTTTCTG 864
DB 945 GACGCTTACAGATTGTGAAAGAAAGAGCGCGCTTCATCTGCGCAACTTCACTTCTG 1004
QY 865 GCGCAACTCTGGAATGAGAAAGATTAGAACCAACATGAGCATCAGGCGCAAG 924
DB 1005 GCGCAGCTGCTGAGTACGAGCGCAGCTGAGCTGCTGCGCTGCAAGGCGAGACCG 1064
QY 925 AGCAAACTCAAGCTGCTGACCTGAGAGCAAAATGAATGAACTGTCTGCTGCTGAG 964
DB 1065 GGCA-----CCCCCTCAGAGAGCGCGGAGCTCCGCGCCAGTCTGCGCGG 1112
```

```

QY 985 GGTGACAGAAAGCCAGACGCGCTCTCACTGCTGAGCCCTGCTGCTGCTGCTGAG 1044
DB 1113 GCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1172
QY 1045 GCAAGCAGACAAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
DB 1173 GC-----CAGGAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220
QY 1105 TCGCTGTAGAGAGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1164
DB 1221 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1280
QY 1165 AGCTGGAAGACAGCAATTAAGCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1219
DB 1281 CCGCTGAGAGACATCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1335
```

RESULT 4

US-09-016-434-91

Sequence 91, Application US/09016434

Patent No. 650938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young

APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

NUMBER OF SEQUENCES: 1490

CORRESPONDENCE ADDRESS:

ADDRESS: INCYTE PHARMACEUTICALS, INC.

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,434

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 91:

SEQUENCE CHARACTERISTICS:

LENGTH: 279 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: LUNGFE703

CLONE: 1234795

US-09-016-434-91

Query Match

11.2%; Score 223; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 1.1e-64;

Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGCGCCATGAGATGTTGAACTCAATTTTACTGAGAGGTTGGTGGCTGCTGAGAA 60
DB 25 ATGCGCCATGAGATGTTGAACTCAATTTTACTGAGAGGTTGGTGGCTGCTGAGAA 84
```

QY 61 AGTGAACGGAAGAGCTGCTAATTGATAGCCGCCATTGGGAAATACATACATCC 120
DB 85 AGTGAACGGAAGAGCTGCTAATTGATAGCCGCCATTGGGAAATACATACATCC 144
QY 121 CACATTTGGAGCCATTATATCACTGCTCCAGCTTATGAGCAAGGTTGCAACAG 180
DB 145 CACATTTGGAGCCATTATATCACTGCTCCAGCTTATGAGCAAGGTTGCAACAG 204
QY 181 GACAAAGTTTAATTACAGAGCTCATCCAGCATTCAGCGAAAC 223
DB 205 GACAAAGTTTAATTACAGAGCTCATCCAGCATTCAGCGAAAC 247

RESULT 5
US-09-557-921-1
; Sequence 1, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125, 416
; CURRENT APPLICATION NUMBER: US/09/557, 921
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-557-921-1

Query Match 6.2%; Score 124.6; DB 4; Length 1830;
Best Local Similarity 56.7%; Pred. No. 7, 2e-31;
Matches 251; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

QY 461 TTGCCAATTTGGGCAACCCGGAATTTTCCCAATCTTTATCTTGGCTCCAGAGATG 520
DB 1036 TCGAGAACGCTGAGCTCACCCCACTTCTTCTTCTTGGCAATGAGCAGATG 1095
QY 521 TCTCAACAAGAGCTGATACAGAGATGGAATGGTATGTTAATGCAAGCTATA 580
DB 1096 CTACAGACCTGGACACCATGACGCGCTGAACATCGCTACGTCATACACATC 1155
QY 581 CCTGTCCAAAGCTGACTTT--ATCCCGAGTCTCATTTCTGCGTGTGCTGATG 637
DB 1156 ATCTTCCCTCTACACTATGAGAAAGGCTGTTCACTACAGCGGCTGCCAGCATG 1215
QY 638 ACAGCTTTTGTGAGAAATTTTGGCGGTGGAGCAAAATCAGTAGATTTTATTGAGAAAG 697
DB 1216 ACAGCAACAAGCAAGAACCTCGGACGATCTTTGAAAGGCTTTTGAATTTGAGGAG 1275
QY 698 CAAGAAGCTCAATGATGATGTTCTAGTGCATGTTTAAAGTGGGATCTCCGCTCCGCCA 757
DB 1276 CTACACAGTGTGGAGAGGGGCTTCTCATCTCACTGCGAGGCTGCGGCTCCGCCA 1335
QY 758 CCAATGCTATCGCTTACATCATGAGAGATGAGCATGCTTTTATGATGAGCTTACAGAT 817
DB 1336 CCAATGCTATCGCTTACATCATGAGAGATGAGCATGAGTATGATGATTAAT 1395
QY 818 TTGTGAAAGAAAAAGCTTATCTTCAAACTTCAATTTTCTGGGCAACTCTG 877
DB 1396 TTGTGAAAGAAAAAGCTTATCTTCAAACTTCAATTTTCTGGGCAACTCTG 1455
QY 878 ACTATGAGAGAGATTAAGAAC 900
DB 1456 AGTTGAGAGAGACTTAACAC 1478

RESULT 6
US-09-922-146-3
; Sequence 3, Application US/09922146

Patent No. 6566133
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; APPLICANT: Brett P. Monia
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION
; FILE REFERENCE: RTS-0252
; CURRENT APPLICATION NUMBER: US/09/922, 146
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 3
; LENGTH: 2303
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (114)...(1268)
US-09-922-146-3

Query Match 5.4%; Score 106.8; DB 4; Length 2303;
Best Local Similarity 55.3%; Pred. No. 8, 7e-25;
Matches 229; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

QY 475 CCAACCCGAATTTCTCCAACTTTATCTTGGCTGCGAGAGATGCTCAACAGAG 534
DB 723 CCGTCCAGATCTGCGCAACCTTATCTGGGAGAGGCGGATTCGGCAATTTGAG 782
QY 535 CTGATACAGCAAGATGGATGTTATGTTTAAATGCCAGCTATACCTGTCAA--G 591
DB 783 AGCCTGCCAATGAGGATCCGCTACATCTCAATGATCACCCCAACCTCCAACTTC 842
QY 592 CCGTACTTTATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTAG 851
DB 843 TTGAGAGAGATGATGATCTTACTTCAAGAGATCCCATCTCCAGACCTGAGCCAG 902
QY 652 AAAATTTCCGCTGTGAGCAAAATGATGATTTTATGAGAAACAAAGCTTCAAT 711
DB 903 AACCTGCGGGTCTTTCGAGGCGCATTTGATGATGAGGCTGTGTCAGAAC 962
QY 712 GATGTTTCTATGACATGTTTATGCTGGATTCCTCGGCTCCGACATGCTATGCC 771
DB 963 TCGGGGCTCTCTCACTGCTTGGGGGCTCAGCGCTTCTGTCACCGTCACTGTGGC 1022
QY 772 TACATGAGAGAGATGAGCATGCTTTAGATGAGATTAAGATTTGTAAGAAAA 831
DB 1023 TACCTATGAGAGATGAGCATGCTTCTCTCAAGATGCTTATACCTGTGCAAGAGAG 1082
QY 832 AGACTTATATCTCCAACTTCAATTTTGGGCAACTCTGTGACTATAG 885
DB 1083 AGCTTATATCTCCCAACTTCACTTATGAGGAGATGCTGTGACTTTGAG 1136

RESULT 7
US-09-016-434-1135
; Sequence 1135, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2109 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1418933
;
US-09-016-434-1135

Query Match
Best Local Similarity 4.8%; Score 96.2; DB 4; Length 2109;
Matches 222; Conservative 0; Mismatches 188; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTATCTGGCTGCGCAGAGATGTCCTCAACAAGAGCTGATACG 543
DB 979 ATCTTCCCTTCTCTACTTGGGCTGCGCAAGACTCCCAACTTGGACGTTGGAG 1038
QY 544 CAGATGGGATGGTATGTTAATGCCAGCTTATACCTGTCCAAGC---CTGACTTT 600
DB 1039 GAATTCGGCATCAAGTACATCTTGAAGTCACCCCAATTTGCCAATCTCTTTAGAAC 1098
QY 601 ATCCCGAGTCTATTTCTGCGCTGCTGCTGAATGACAGCTTTGTGAGAAATTTTG 660
DB 1099 GCAGAGAGTTTAATCAAGCAAAATCCCATCTCGATCACTGAGCCAAACCTGTCC 1158
QY 661 CCGTGTGGCAAAATCAGTATGATTTGATGAAAGAAAGCCCTCCATGATGATGTT 720
DB 1159 CAGTTTTTCCCTGAGCCATTTCTTTCATGATGAAAGCCCGGGCAAGAACGTGGTGC 1218
QY 721 CTAGTGACCTGTTAGCTGGATCTCCGCTCCGCAACATGCGTATGCTTACATATG 780
DB 1219 TTGTATCATTTGCTGGCTGGCATTAGCCGCTCAGTCACTGATGCTGTACCTTATG 1278
QY 781 AAGAGATGAGCATGCTTTAGATGAGCTTACAGATTTGTGAAAAAGAAAGACTACT 840
DB 1279 CAGAACTCAATCTGTCATGATGAGATGCTATGACATTTGCAAAATGAAAAATCCAAC 1338
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCCGAGCATATGAGAAAGAT 893
DB 1339 ATATCCCTCAACTCACTTCAATGGGTCACTGTGAGCTTGGAGAGAGCT 1391

RESULT 8
US-09-023-655-946
; Sequence 946 Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
```

```

;
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: US/09/023,655
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 946:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2109 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1418933
;
US-09-023-655-946

Query Match
Best Local Similarity 4.8%; Score 96.2; DB 4; Length 2109;
Matches 222; Conservative 0; Mismatches 188; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTATCTGGCTGCGCAGAGATGTCCTCAACAAGAGCTGATACG 543
DB 979 ATCTTCCCTTCTCTACTTGGGCTGCGCAAGACTCCCAACTTGGACGTTGGAG 1038
QY 544 CAGATGGGATGGTATGTTAATGCCAGCTTATACCTGTCCAAGC---CTGACTTT 600
DB 1039 GAATTCGGCATCAAGTACATCTTGAAGTCACCCCAATTTGCCAATCTCTTTAGAAC 1098
QY 601 ATCCCGAGTCTATTTCTGCGCTGCTGCTGAATGACAGCTTTGTGAGAAATTTTG 660
DB 1099 GCAGAGAGTTTAATCAAGCAAAATCCCATCTCGATCACTGAGCCAAACCTGTCC 1158
QY 661 CCGTGTGGCAAAATCAGTATGATTTGATGAAAGAAAGCCCTCCATGATGATGTT 720
DB 1159 CAGTTTTTCCCTGAGCCATTTCTTTCATGATGAAAGCCCGGGCAAGAACGTGGTGC 1218
QY 721 CTAGTGACCTGTTAGCTGGATCTCCGCTCCGCAACATGCGTATGCTTACATATG 780
DB 1219 TTGTATCATTTGCTGGCTGGCATTAGCCGCTCAGTCACTGATGCTGTACCTTATG 1278
QY 781 AAGAGATGAGCATGCTTTAGATGAGCTTACAGATTTGTGAAAAAGAAAGACTACT 840
DB 1279 CAGAACTCAATCTGTCATGATGAGATGCTATGACATTTGCAAAATGAAAAATCCAAC 1338
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCCGAGCATATGAGAAAGAT 893
DB 1339 ATATCCCTCAACTCACTTCAATGGGTCACTGTGAGCTTGGAGAGAGCT 1391

RESULT 9
US-09-023-655-347
; Sequence 347 Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
```

APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 347:
SEQUENCE CHARACTERISTICS:
LENGTH: 1208 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THYRNOT03
CLONE: 1444245
US-09-023-655-347

Query Match 4.7%; Score 94.6; DB 4; Length 1208;
Best Local Similarity 64.3%; Pred. No. 6.7e-21;
Matches 142; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 680 TAGATTTCATTGGAAGCAAAAGCCCTCAATGAGTGTCTAGTGCACGTTTGGCTG 739
DB 4 TTGATTCATTGGAAGCAAGCTCAACAGTGTGGAAAGGGGCTTTCATCATCGCAGGCTG 63
QY 740 GGATTCCTCCCTCGCCACCATGCTATCGCTACATCATGAGAGAGATGATGCTT 799
DB 64 GGGTGTCCGCTCGCCACCATGCTATCGCTTACTTATGATGAGCACTCGATGCA 123
QY 800 TAGATGAAGCTTACAGTTTGTGAAAGAAAAAGACTTACTATATCTCCAACTTCAAT 859
DB 124 TGATGATGCTTATTAATTTGTCAAAAGCAAGCAACCAATTAATCTCCCAACTTAAT 183
QY 860 TTCTGGGCGCAACCTCGGACTATGAGAAAGATTGAAGC 900
DB 184 TCATGGGCGAGTGTGAGTTGAGGAGAACTTAACAC 224

RESULT 10
US-09-702-705-801
Sequence 801, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Derrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapien
US-09-702-705-801

Query Match 4.7%; Score 93.6; DB 4; Length 1619;
Best Local Similarity 52.3%; Pred. No. 1.8e-20;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 484 ATTCTTCCCAATCTTATCTGCTCCAGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTCTACCTCGGCACTGCTACATGCTGCCGAGAGACATGCTGAC 379
QY 544 CAGATGGATGCTTATGTTAATGCCAGCTATACCTGTCCAAAGCTGACTTTATC 603
DB 380 GCCCTGGGATCACGGCTGTGTAATGTCCTCGGACTGCCCAACCC---ACTTGAA 436
QY 604 CCGAGTCTATTCCTGGGTGCGTGAATGACAGCTTTGTGAAATTTGGCCG 663
DB 437 GGACACTATGATAGTAGTCAATCCAGTGAAGATTAACAAGGCCGACATCAGCTCC 436
QY 664 TGGTTGACAAATCAGTATGTTTCAATGAGAAAGCAAAAGCCCTCAATGATGTTCTA 723
DB 497 TGGTTATGGAAGCCATAGATCATCATGATCGGTGAAGACGCGGTGGCGCGTGTG 556
QY 724 GTGCACTGTTAGCTGGGATCTCCGCTCGCCACCATGCTATCGCTTACATGAG 783
DB 557 GTGCACTGCGAGCGGAGCATCTCGCGTGGCCACCATCTGCTGCTACCTGATGATG 616
QY 784 AGATGACATGCTTTAGATGAAGCTTACAGTTTGTGAAGAAAAAGACTTACTATA 843
DB 617 AAGAAACGGGTGAGCTGAGAGAGGCTTGAAGTTGTTAAGCAGCGCCGACATTAATC 676
QY 844 TTCCAACTTCAATTTCTGGCCCACTCTGGAATGAGAAAGAAATTGAAGACAG 903
DB 677 TGCCCACTTCACTTATGAGGAGCTGCTGCAATGAGTCCAGGTGCTGGCAGC 736
QY 904 ACTGAGCATCAGGGCCAAAGAGC 927
DB 737 TCCTGTGTGCGGAGGCTGCTAGC 760

RESULT 11
US-09-736-457-801
Sequence 801, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Derrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, LiJun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-801

Query Match 4.7%; Score 93.6; DB 4; Length 1619;
Best Local Similarity 52.3%; Pred. No. 1.8e-20;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTTATCTTGGCTGCCAGGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTCTTACCTCGGAGTGCCTTACCATGCTCCGAGAGACATGCTGAC 379
QY 544 CAGAAATGGATTTGGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGGCAATCAGGCTCTGTTGAATGTTCTCTCGAAGCTGCCAAAC--ACTTTGAA 436
QY 604 CCCGAGTCTATTTCTGCGGTGCTGCTGATGACCTTTGTGAGAAATTTTCCG 663
DB 437 GACACTATCAGTACAAAGTGCATCCAGTGAAGTAAACAAGGCCGACATCAGCTC 496
QY 664 TGGTTGACAAATTCAGTATTTCTTGAAGAAAGCAAGCTCCAAATGATGTTCTTA 723
DB 497 TGGTTATGAAAGCAATAGATGATCATGCTGTAAGAGATGCTCGTGGGCCGTGCTG 556
QY 724 GTGCACTGTTAGCTGGATCTCCGCTCGGCAACCTGCTATGCTTACATGATGAAG 783
DB 557 GTGCACTGCAAGGCGGGATCTCGGGTGGCCACCATCTGCTGCTTACCTGATGATG 616
QY 784 AGATGACATGCTTTTGAAGTGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACTATA 843
DB 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTCAAGTGTGAAGAGCGCGGACGACTATC 676
QY 844 TCTCCAACTTCAATTTCTGGGCCAATCTCTGACCTATGAGAGAAATTAAGAACAG 903
DB 677 TCGCCCACTTCAAGCTTATGAGGAGCTGCTCTCAAGTTCAGATCCAGTGTGCTGACAG 736
QY 904 ACTGAGCATCAGGGCCAAAGAGC 927
DB 737 TCTGTGCTGCGGAGGCTGCTAGC 760

RESULT 12

US-09-614-124B-801
Sequence 801, Application US/09614124B
Patent No. 6630574
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.478C9
CURRENT APPLICATION NUMBER: US/09/614,124B
CURRENT FILING DATE: 2001-07-11
NUMBER OF SEQ ID NOS: 1668
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapien
US-09-614-124B-801

Query Match 4.7%; Score 93.6; DB 4; Length 1619;
Best Local Similarity 52.3%; Pred. No. 1.8e-20;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTTATCTTGGCTGCCAGGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTCTTACCTCGGAGTGCCTTACCATGCTCCGAGAGACATGCTGAC 379
QY 544 CAGAAATGGATTTGGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGGCAATCAGGCTCTGTTGAATGTTCTCTCGAAGCTGCCAAAC--ACTTTGAA 436
QY 604 CCCGAGTCTATTTCTGCGGTGCTGCTGATGACCTTTGTGAGAAATTTTCCG 663
DB 437 GACACTATCAGTACAAAGTGCATCCAGTGAAGTAAACAAGGCCGACATCAGCTC 496
QY 664 TGGTTGACAAATTCAGTATTTCTTGAAGAAAGCAAGCTCCAAATGATGTTCTTA 723
DB 497 TGGTTATGAAAGCAATAGATGATCATGCTGTAAGAGATGCTCGTGGGCCGTGCTG 556
QY 724 GTGCACTGTTAGCTGGATCTCCGCTCGGCAACCTGCTATGCTTACATGATGAAG 783
DB 557 GTGCACTGCAAGGCGGGATCTCGGGTGGCCACCATCTGCTGCTTACCTGATGATG 616
QY 784 AGATGACATGCTTTTGAAGTGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACTATA 843
DB 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTCAAGTGTGAAGAGCGCGGACGACTATC 676
QY 844 TCTCCAACTTCAATTTCTGGGCCAATCTCTGACCTATGAGAGAAATTAAGAACAG 903
DB 677 TCGCCCACTTCAAGCTTATGAGGAGCTGCTCTCAAGTTCAGATCCAGTGTGCTGACAG 736
QY 904 ACTGAGCATCAGGGCCAAAGAGC 927
DB 737 TCTGTGCTGCGGAGGCTGCTAGC 760

RESULT 13

US-09-671-325-801
Sequence 801, Application US/09671325
Patent No. 6667154
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C12
CURRENT APPLICATION NUMBER: US/09/671,325
CURRENT FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 1825
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapien
US-09-671-325-801

Query Match 4.7%; Score 93.6; DB 4; Length 1619;
Best Local Similarity 52.3%; Pred. No. 1.8e-20;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTTATCTTGGCTGCCAGGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTCTTACCTCGGAGTGCCTTACCATGCTCCGAGAGACATGCTGAC 379

QY 544 CAGATGGAGTGTATGTTAATGACGCTATACCTGTCGAAGCCTGACTTATC 603
DB 380 GCCCTGGGATCAGCGGCTGTGTAATGTCTCTCGAGCTGCCAACC---ACTTTGAA 436
QY 604 CCGAGTCTCATTTCTCGCTGTGCTGTGAATGACAGCTTTGTGAAATTTTCCG 663
DB 437 GGAACATATAGTAAAGTGCATCCGAGTGAAGATTAACCAAGGCCAGCATCACTCC 496
QY 664 TGGTTGACAAATCAGTAAATTTTCAATGAAAAGCAAAACCTTCATGATGTCTTA 723
DB 497 TGGTTACATGAAACCAATGAGTACATGATGCGGTGAAGAGATGCGCTGGCCGCTG 556
QY 724 GTGCACTGTTAGTGGGATCTCCGCTCCGCCACATCGCTATGCGCTCATATGAAG 783
DB 557 GTGCACTGCCAGCGGCGGATCTCGCGGTCCGCCACATCTGCTGCTGCTGATGATG 616
QY 784 AGGATGACATGCTTTAGTAAAGCTTACAGATTTGTGAAAAGAAAAAGACCTACTATA 843
DB 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTCGAGTTGTTAAGACGCGCCAGCATTTATC 676
QY 844 TCTCCAACTTCAATTTTCTGGGCCCACTCTGACTATAGAAAGATTAAGAACAG 903
DB 677 TCGCCCAACTTCACTTCAATGAGGAGCTGCTGCAATTCAGATCCAGGTGCTGCCACG 736
QY 904 ACTGAGCATCAGGGGCCAAAGAGC 927
DB 737 TCTGTGCTGCGGAGGCTGCTAGC 760

RESULT 14

US-09-589-184-801
; Sequence 801, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaltanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR LUNG CANCER
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589.184
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-589-184-801

Query Match 4.7%; Score 93.6; DB 4; Length 1619;
Best Local Similarity 52.3%; Pred. No. 1.8e-20;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTATCTTGGCTGCCAGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTCTCTACTGAGGAGTGTCTACATGCTGCCGAGAGACATGCTGAC 379
QY 544 CAGATGGAGTGTATGTTAATGACGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGGATCAGCGGCTGTGTAATGTCTCTCGAGCTGCCAACC---ACTTTGAA 436
QY 604 CCGAGTCTCATTTCTCGCTGTGCTGTGAATGACAGCTTTGTGAAATTTTCCG 663
DB 437 GGAACATATAGTAAAGTGCATCCGAGTGAAGATTAACCAAGGCCAGCATCACTCC 496
QY 664 TGGTTGACAAATCAGTAAATTTTCAATGAAAAGCAAAACCTTCATGATGTCTTA 723
DB 497 TGGTTACATGAAACCAATGAGTACATGATGCGGTGAAGAGATGCGCTGGCCGCTG 556
QY 724 GTGCACTGTTAGTGGGATCTCCGCTCCGCCACATCGCTATGCGCTCATATGAAG 783
DB 557 GTGCACTGCCAGCGGCGGATCTCGCGGTCCGCCACATCTGCTGCTGCTGATGATG 616
QY 784 AGGATGACATGCTTTAGTAAAGCTTACAGATTTGTGAAAAGAAAAAGACCTACTATA 843
DB 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTCGAGTTGTTAAGACGCGCCAGCATTTATC 676

DB 497 TGGTTACATGAAACCAATGAGTACATGATGCCGTGAAGAGATGCCGTGGCCGCTGCTG 556
QY 724 GTGCACTGTTAGTGGATCTTCCGCTCCGCCACATGCTATGCTTATCATGAAAG 783
DB 557 GTGCACTGCCAGCGGCGGATCTCGCGGTCCGCCACATCTGCTGCTGCTGATGATG 616
QY 784 AGGATGACATGCTTTAGTAAAGCTTACAGATTTGTGAAAAGAAAAAGACCTACTATA 843
DB 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTCGAGTTGTTAAGACGCGCCAGCATTTATC 676
QY 844 TCTCCAACTTCAATTTTCTGGGCCCACTCTGACTATAGAAAGATTAAGAACAG 903
DB 677 TCGCCCAACTTCACTTCAATGAGGAGCTGCTGCAATTCAGATCCAGGTGCTGCCACG 736
QY 904 ACTGAGCATCAGGGGCCAAAGAGC 927
DB 737 TCTGTGCTGCGGAGGCTGCTAGC 760

RESULT 15

US-09-702-705-804
; Sequence 804, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaltanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702.705
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 804
; LENGTH: 4637
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-804

Query Match 4.7%; Score 93.6; DB 4; Length 4637;
Best Local Similarity 52.3%; Pred. No. 4.3e-20;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTATCTTGGCTGCCAGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTCTCTACTGAGGAGTGTCTACATGCTGCCGAGAGACATGCTGAC 379
QY 544 CAGATGGAGTGTATGTTAATGACGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGGATCAGCGGCTGTGTAATGTCTCTCGAGCTGCCAACC---ACTTTGAA 436
QY 604 CCGAGTCTCATTTCTCGCTGTGCTGTGAATGACAGCTTTGTGAAAATTTTCCG 663
DB 437 GGAACATATAGTAAAGTGCATCCGAGTGAAGATTAACCAAGGCCAGCATCACTCC 496
QY 497 TGGTTACATGAAACCAATGAGTACATGATGCGGTGAAGAGATGCGCTGGCCGCTG 556
QY 724 GTGCACTGTTAGTGGGATCTCCGCTCCGCCACATCGCTATGCGCTCATATGAAG 783
DB 557 GTGCACTGCCAGCGGCGGATCTCGCGGTCCGCCACATCTGCTGCTGCTGATGATG 616
QY 784 AGGATGACATGCTTTAGTAAAGCTTACAGATTTGTGAAAAGAAAAAGACCTACTATA 843
DB 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTCGAGTTGTTAAGACGCGCCAGCATTTATC 676

Oy 844 TCTCCAAACTTCAATTTCTGGGCAACTCTGACTATGAGAAGAATTAAACAACAG 903
 Db 677 TCGCCAACTTCACTTCAATGGGGCAAGTCTGAGTCCAGGTGCTGGCAAG 736
 Oy 904 ACTGAGCATCAGGGCCAAAGAGC 927
 Db 737 TCTGTGTGCTGCGAGGCTGCTAGC 760

Search completed: June 22, 2004, 02:56:08
 Job time : 168.047 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:21 ; Search time 95.5895 Seconds
(without alignments)
11793.817 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532
Predicted score: 3615
Sequence: 1 atggcccatgagatgattg.....tggaaatcatgaggtctcc 1995

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=xjp
-Q/cgnt 1/USPTO.spool.p/US10029345/runat.21062004.122814.4167/app.query.fasta.1.3278
-DB=A.GeneSeq.29Jan04 -OPMT=fastaan -SUFFIX=rag -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -INITs=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US10029345 -CCGN 1.128 -runat.21062004.122814.4167 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: A_GeneSeq.29Jan04:.*
1: geneSeq1980s:.*
2: geneSeq1990s:.*
3: geneSeq2000s:.*
4: geneSeq2001s:.*
5: geneSeq2002s:.*
6: geneSeq2003as:.*
7: geneSeq2003bs:.*
8: geneSeq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	3418	94.6	665 5	ABR52381 Protein x
2	3418	94.6	665 5	ABR52407 Protein x
3	3406	94.2	665 4	AAE04834 Human SGP
4	3406	94.2	665 4	AAU09016 Human dua
5	3406	94.2	665 5	AAU09156 Human dua
6	3406	94.2	665 5	AAU09946 Protein s
7	3406	94.2	665 5	AAU75789 Human pro
8	3406	94.2	665 5	ABR97946 Human pro
9	3406	94.2	665 5	AAU79929 Human dua
10	3406	94.2	665 5	ABR97291 Novel hum

11	3399	94.0	665 6	ADA54744 Human pro
12	3398	94.0	665 5	AAU79161 Human dua
13	3396	93.9	665 5	AAU79162 Human dua
14	3388	93.7	665 5	ABR52352 Protein x
15	3379.5	93.5	666 4	ABR520325 Human pro
16	3368.5	93.2	664 5	ABR52424 Protein x
17	3322	91.9	672 4	AAU52744 Human pro
18	3057.5	84.6	660 5	ABR52385 Protein x
19	2594	71.8	517 5	AAU79159 Human dua
20	1542	42.7	302 5	ABR52425 Protein x
21	1326	36.7	625 5	ABR52382 Protein x
22	1326	36.7	625 5	ABR52350 Protein x
23	1326	36.7	625 6	ABG73440 Human dua
24	1302	36.0	663 5	ABR52351 Protein x
25	1297	35.9	663 2	AAU79150 Protein x
26	977.5	27.0	579 7	ADP08458 Novel pro
27	742.5	20.5	375 4	ABG07902 Novel hum
28	741	20.5	140 5	ABR52404 Peptide x
29	716	19.8	140 5	ABR52405 Peptide x
30	654.5	18.1	170 4	AAE66436 Human MAP
31	654.5	18.1	170 4	AAE06780 Human dua
32	476	13.2	482 4	ABR73225 Human pro
33	471.5	13.0	381 5	ABR52383 Protein x
34	471.5	13.0	381 6	ABP96803 Human COP
35	469	13.0	444 5	AAO20515 Protein o
36	469	13.0	482 3	ABR18655 A human r
37	469	13.0	482 3	ABR29641 Human dua
38	469	13.0	482 5	ABR52384 Protein x
39	469	13.0	482 5	ABJ05600 Breast ca
40	466.5	12.9	381 6	ABP57087 Mouse MKP
41	465.5	12.9	381 7	ADD48300 Rat Prote
42	465.5	12.9	381 7	ADP62625 Rat Prote
43	455	12.6	394 4	ABR76875 Human lun
44	455	12.6	394 5	AAU85530 ST18 lung
45	455	12.6	394 6	ABU69502 Humna STY

ALIGNMENTS

RESULT 1	ID	ABR52381 standard; protein; 665 AA.
AC	ABR52381;	
XX		
DT	19-JUN-2003 (first entry)	
XX		
DE	Protein relating to the invention SEQ ID NO: 109.	
XX		
KW	antiproliferative; hepatotropic; nephrotropic; antiarthritic;	
KW	antiproliferative; cardiatic; cytosolic; gene therapy; liver disease;	
KW	proliferative disorder; renal failure; cardiovascular disorder;	
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;	
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200257460-A2.	
XX		
PD	25-JUL-2002.	
XX		
PF	20-DEC-2001; 2001MO-US050459.	
XX		
PR	20-DEC-2001; 2000US-0256668P.	
PR	30-MAR-2001; 2001US-0280186P.	
PR	01-MAY-2001; 2001US-0287735P.	
PR	05-JUN-2001; 2001US-0295488P.	
PR	25-JUN-2001; 2001US-0300465P.	
XX		
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.	
XX		
PI	Jackson DG, Feder J, Nelson T, Muntier G, Ramanathan C, Lee L, Stiemens N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;	

PI Krystek S, Mcatee P, Suchard S, Banas D,
XX
DR MPI; 2002-599721/64.
XX N-PSDB; ACC60559.
PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
XX disorders.
XX
XX Claim 5; Fig 12; 801pp; English.
CC The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
XX sequence is used in the exemplification of the invention
SQ Sequence 665 AA;

Alignment Scores:
Pred. No.: 2,85e-288 Length: 665
Score: 3418.00 Matches: 665
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.55% Indels: 0
DB: 5 Gaps: 0

US-10-029-345A-108_COPY_538_2532 (1-1995) x ABR52381 (1-665)
QY 1 ATGGCCCATGAGATGATGTAAGTCAATATTGATCGAGAGGTTGGCTCTGCTGAA 60
DB 1 MetalaHISGiumetilegLYThnGlnileValThnGlnrgeuValAlaleuendGlu 20
QY 61 AGTGAACGGAAGAAAGGCTGCTAATTGATGAGCCGCAATTTGGAAATACATCATCC 120
DB 21 SerGlythrGluysValleuLeuileApsSerArGProPheValGluTyAsnthSer 40
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAAACGGAAGTTGCAACG 180
DB 41 HisleleuGluAlaleuAenleAenySerlySleuMeLySArGArGleuGln 60
QY 181 GACAAAGTGTATTAACAGGCTCATCCGCAATTCAGGAAACATTAAGTTGACATTGAT 240
DB 61 AsplyValleuileThnGlnleuileGlnHisSerAlaLysHLSlyValAaplleAps 80
QY 241 TGCAGTCAGAAAGGTTGATTTACGATCAAGCTCCCAAGATGTTGGCTCTCTCTCA 300
DB 81 CysSerGlnLysValValValTyArGpInserSerGlnApsValAlaSerleuSerSer 100
QY 301 GACTGTTTCTCACTGTTCTTGCTGTAATCTGGAAGAAGCTTCACTGTTCACTG 360
DB 101 AspCysPheleuThValleuLeuGlyLysleuGlnLysSerPheAenSerValHisleu 120
QY 361 CTTCGAGGTGGGTTTGCTGAGTCTCTCGTGTGTTCCCTGCTCTGTAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArGysPheProGlyLeuGlnGlyLysSer 140
QY 421 ACTTAGTCCCTACCTGATTTCTCAGCTTGGCTTCACTGTTGCCAACTTTGGGCAAC 480
DB 141 ThrleuValProthrCysHisSerGlnProCysleuProValAlaAsnlelyProthr 160
QY 481 CGAATTTCTCCCAATCTTTATCTTGCTGCTGCGAGCGAGATGCTCTCAAGAAGCTGATA 540
DB 161 ArgileuProAsnleuTyrlleuGlyCysGlnApsValleuAsnLysleuile 180

QY 541 CAGCAAGATGGGATTTGTTATGTTTAAATGCCATTAACCTGTCCAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyLylelyTyValleuAsnHisSerlyrThrCysProLysProApsPhe 200
QY 601 ATCCCGAAGTTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 201 IleProGlnSerHisPheleuAryValProValAsnApsSerPheCysGlnLysleu 220
QY 661 CCGTGGTTCGAAATCAATGATGATTTCAATTGAGAAACCAAAAGCTTCATGATGTT 720
DB 221 ProThrleuApsLysSerValApsPheileuLysAlaLysAlaSerleuGlyVal 240
QY 721 CTAGTCACCTGTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 241 LeuValHisCysleuAlaGlyIleSerArGSerHisThleAlleAlaTyrlleMet 260
QY 781 AAGAGATGAGCATGCTTATGATGAACTTACAGATTTGTGAAAGAAAAAGACTACT 840
DB 261 LysArGMetApsMetSerleuApsGlnAlaTyArGpHeVallyGlyLysArProthr 280
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCCAATCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 281 IleserProAsnApsPheAsnApsPheleuGlyLysleuApsTyrlleLysAsn 300
QY 901 CAGACTGAGCATCAGAGGCGCAAGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 301 GlnThrGlyAlaSerGlyProLysSerLysLysleuHisleuGlnLysProAsn 320
QY 961 GAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 321 GlnProValProAlaValSerGlnGlyGlyGlnLysSerGlnThrProLysProPro 340
QY 1021 TGTGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 341 CysAlaApsSerAlaThnSerGlnAlaAlaGlyGlnApsProValHisProAlaSerVal 360
QY 1081 CCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 361 ProSerValProSerValGlnProSerleuGlnApsSerProleuValGlnAlaLeu 380
QY 1141 AGTGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 381 SerGlyleuHisleuSerAlaApsArGleuGlnApsSerAsnLysleuLysArSerPhe 400
QY 1201 TCTCGATTAATCAATGATTTGATATTCATGAGCCAGATGAGCATCTTACATGCTTC 1260
DB 401 SerleuApslleLysSerValSerTyrlSerAlaSerMetAlaSerleuHislelyPhe 420
QY 1261 TCTCATCAAGATGCTTGGAAATCTCAAACTTCACTGCTGATGGAGCAAC 1320
DB 421 SerSerSerGlnApsAlaLeuGlnLysTyrlLysProSerThrThnleuApsGlyThnAsn 440
QY 1321 AAGCTAAGCCAGTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
DB 441 LysleuGlnGlnApsPheSerValGlnGlnleuSerGlnGlnThrProGlnThrSerPro 460
QY 1381 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGCAAGCCGCGAGCTTCAAGACG 1440
DB 461 AsplyGlnGlnAlaSerleleProLysLysleuGlnThrAlaApsProSerApsSerGln 480
QY 1441 AGCAAGCATTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 481 SerLysArGleuHisSerValArGThnSerSerSerlelyThrAlaGlnApsSerleu 500
QY 1501 TCTCCACTGATTCGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
DB 501 SerProleuHisArGSerGlySerValGlnApsAsnLysThnSerPheleuPheGly 520
QY 1561 CTTTCACACACGACGACCTTCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB 521 leuSerThrSerGlnGlnHisleuThrLysSerAlaGlyleuGlyleuLyslelyTyrlHis 540

Db 221 ProTrpLeuAspLysSerValAspHe11eGluValAlaLysSerAengLyCysVal 240
 QY 721 CTAGTGCACCTTTTACCTGGGATCTCCGCTCCGACCAATCGCTATCGCTTCATCATG 780
 Db 241 LeuValHisCysLeuAlaGly11eSerArSerSerAlaThrLeuAlaLeuAla1r11eMet 260
 QY 781 AAGAGATGACATGCTTTTATGATGAGCTTACGATTTGTGAAAGAAAAAGAACTACT 840
 Db 261 LysArgMetAspMetSerLeuAspGluAlaTyArpHeVal11uLysArgProThr 280
 QY 841 ATATCTCAACTTCAATTTTCTGGCCAACTCTGACATGAGAGAAAGATTAGAAC 900
 Db 281 IleSerProMetPheAspPheLeuGlyGlnLeuLeuAspArgGluLysLysLysLys 300
 QY 901 CAGACTGGAGCATGAGGCGCAAGAGCAAACTCAAGCTGCTGACCTGGAGAAAGCAAT 960
 Db 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHis11eGluLysProAsn 320
 QY 961 GAACCTGTCCCTGTCTCTCAGAGGGGTGACAGAAAAAGCGAACCCCTCAGTCCACCC 1020
 Db 321 GluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProPro 340
 QY 1021 TGTGCCACTCTCTCTACCTCAGAGGACAGAGCAAAAGCCCTGACATCCCGCAAGCTG 1080
 Db 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 QY 1081 CCCAGGCGTCCGCGCTGACGCGCTGCTTACAGAGCAAGCCCGCTGTACAGAGCGCTC 1140
 Db 361 ProSerValProSerValGlnProSerLeuGluLysSerProLeuValGlnAlaLeu 380
 QY 1141 AGTGGCTGCACCTCTCCGACAGACAGCTGGAGAGCAAAATAGCTCAAGCGTCTTC 1200
 Db 381 SerGlyLeuHis11eSerAlaAspArgLeuGluLysSerAsnLysLeuLysArgSerPhe 400
 QY 1201 TCTCTGATATGAAATTCATTTTATTCAGCCAGCATGGCAGCATCTTACATGAGCTTC 1260
 Db 401 SerLeuAspLysLeuSerValSerLysSerAlaSerMetAla11eSerLeuHisGlyPhe 420
 QY 1261 TCTCTATGAGAAATGCTTTGGAATATCTACAAACCTTCCATCTGATGGAGCAAC 1320
 Db 421 SerSerSerGluAspAlaLeuGluTyTyLysProSerThrThreLeuAsp11yThrAsn 440
 QY 1321 AAGCTATGCCAGTCTTCCCTGTTTCAGAACTATCGAGACAGATCCCGAAACAGTCTCT 1380
 Db 441 LysLeuGlySerGlnPheSerProValGlnGluLeuSerGlnGlnThrProGluThrSerPro 460
 QY 1381 GATATGAGAGAGAGCCAGCATCCCAAGAAAGCTGACAGACCGCAGGCTTTCAGACCCAG 1440
 Db 461 AspLysGluGluAlaSerLysLeuProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 QY 1441 AGCAAGGATTTGATTGCGTCAGAACCCAGACAGATGCGACCCCGCCAGAGGTCCTTTTA 1500
 Db 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 QY 1501 TCTTCCACTGATCGAAGTGGAGCGTGGAGCAATTCACACACACCTTCTTTTGGC 1560
 Db 501 SerProLeuHisArgSerGlySerValGluLysAsnTyHis11eSerPheLeuPheGly 520
 QY 1561 CTTTCCACGACGACAGACACTCAGCAAGTCTGCTGAGCGCTTAAAGGCTGGAC 1620
 Db 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuGlyTyPheHis 540
 QY 1621 TCGGATATCTTGGCCCCCAGACCTTAACCTTCAACCTTCTGACGACGACGCTGATTTTGGC 1680
 Db 541 SerAspLysLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrPyrPheAla 560
 QY 1681 ACAGAGTCTCAGACTTCTACTGCTGCTCAGCCATTCACGAGGACGCTCCAGTTACTCT 1740
 Db 561 ThrGlnSerSerHisPheTyTserAlaSerAla11eTyGlyGlySerAlaSerLysSer 580
 QY 1741 GCCATACGCTGACGACGCTGCCACTTGGCAGAGCAAGTCTAATTTCTGTGGCAGGCGG 1800
 Db 581 AlaTySerCysSerGlnLeuProThrCysGlyAspGlnValTySerValArgArg 600

QY 1801 CAGAAAGCAAGTACAGAGCTGATCTCCGCGGAGACTGCTGCATGAGAGAGCCCTTTGAA 1860
 Db 601 GlnLysProSerAspArgAlaAspSerArgArgSerThrPhe11eGluGlnSerProPheGlu 620
 QY 1861 AACCACTTTTAAACGAGAAAGCTGCCAAATGGAATTTGGAGAGACATCATGTACAGAAC 1920
 Db 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlnSerLysMetSerGluAsn 640
 QY 1921 AGGTCAAGGAGAGCTGGGAGAAAGTGGGACGATGATCTTATGCTTTTGGCAGACATGAA 1980
 Db 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerPheSerGlySerMetGlu 660
 QY 1981 ATCATTTGAGCTCTCC 1995
 Db 661 Ile11eGluValSer 665
 RESULT 3
 AAE04834
 ID AAE04834 standard; protein; 665 AA.
 AC AAE04834;
 XX
 DT 10-SEP-2001 (first entry)
 XX
 DE Human SGP002 phosphatase polypeptide.
 XX
 KW Human; SGP002 phosphatase polypeptide; phosphatase-related disease;
 KW immun-related disorder; organ transplant rejection;
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
 KW metabolic disorder; hematopoietic cancer; mood disorder; cardiac;
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
 KW attention disorder; cognition disorder; psychotic disorder; cytostatic;
 KW neurological disorder; vincristine; neurotropic; cerebroprotective; therapy;
 KW neuroprotective; antibacterial; vulnery; tranquilliser; antihistaminic;
 KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
 KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
 KW MRP; mirafaine; chromosome 12p11.1-p12.1.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..173
 FT /label= Catalytic_domain
 FT Domain 158..297
 FT /label= Phosphatase_domain
 XX
 PN MO200146394-A2.
 XX
 FD 28-JUN-2001.
 XX
 PF 21-DEC-2000; 2000WO-US034736.
 XX
 PR 21-DEC-1999; 99US-0173255P.
 PR 28-DEC-1999; 99US-0175766P.
 PR 25-JAN-2000; 2000US-0178078P.
 PR 31-JAN-2000; 2000US-0179301P.
 XX
 PA (SUBE-) SUGEN INC.
 XX
 PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;
 PI Hill RJ, Flanagan P;
 PI
 XX
 DR WPI; 2001-418058/44.
 DR N-PSDB; AAD09492.
 XX
 PT Novel phosphatase polypeptide useful for treating cancers, immune-related
 PT diseases and disorders, cardiovascular disease, brain or neuronal-
 PT associated diseases and metabolic disorders.
 XX
 PS Claim 7; Fig 2; 186pp; English.
 XX

CC The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorder, including cancers of tissues, cancers of
CC haematopoietic origin, diseases of central and peripheral nervous system,
CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
CC lateral sclerosis, viral infections, infections caused by prions,
CC bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual
CC dysfunction, mood disorders, attention disorders, cognition disorders,
CC hypofunction, hypertension, psychotic disorders, neurological disorders,
CC dyskinesias and organ transplant rejection. The present amino acid
CC sequence is human SGP002 phosphatase polypeptide. This sequence is
CC classified as dual specificity phosphatase (DSP) and Map kinase
CC phosphatase (MKP). SGP002 gene maps to chromosomal position 12p11.1-p12.1
CC
XX
SQ Sequence 665 AA:

Alignment Scores:
Pred. No.: 3.16e-287 Length: 665
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 94.22% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-108_COPY_538_2532 (1-1995) x AAE04834 (1-665)
QY 1 ATGGCCCATGAGATGATGGAGCTCAATGTTACTGAGAGGTTGGTGGCTGTGGAA 60
Db 1 MetAlaHisGlnUetIleGlyIrnGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
QY 61 AGTGAACGGAAGAGTGTCTTAATTTGATAGCCGACATTTGTGAAATCAATACATCC 120
Db 21 SerGlyThrGlnIuValLeuLeuIleAspSerArgProHeValGlnIuYrAanThrSer 40
QY 121 CACATTTTGGAGCCATTATATCACTGCTCCAGCTTATGAAGGAGAGTTGCAACAG 180
Db 41 HisIleLeuGlnIuAlaIleAenIleAsnCySerIuSerIuMetIuSerArgIuGln 60
QY 181 GACAAAGTGTATATTCAGAGCTCATCCAGCATTCAGCGAAACATAGTGTGACATTTGAT 240
Db 61 AspIuValLeuIleThrGlnLeuIleGlnHisSerAlaIuShiIuValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATTTACGATCAAGCTCCCAAGATGTGCTCTCTCTTCA 300
Db 81 CySerGlnIuValValValIuYrAspGlnSerSerGlnAspValAlaSerIuSerSer 100
QY 301 GACTGTTTCTCACTGTACTTCTGGGGTAACTGAGAGAGCTTCAACTGTCTTCACTTG 360
Db 101 AspCyPheLeuThrValLeuLeuGlyIuYrLeuGlnIuYrSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGGTGGGTTTGCTGAGTCTCTCGTTGTTTCCCTGGCTGTGAGAGGAAATCC 420
Db 121 LeuAlaGlyIuYrPheAlaGlnIuPheSerArgCyPheProIuYeuCyGlnIuYrSer 140
QY 421 ACTCTAGTCCCTACCTGCACTTCTCAGCTTGCTTCACTGTGCTCAACATTTGGCCAAAC 480
Db 141 ThrLeuValProThrCyHisSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGACAGAGATGCTTCAACAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuIuYrLeuGlnIuYrGlnArgAspValLeuAsnIuYrGlnLeuMet 180
QY 541 CAGCAGATGGAGATTGGTATGTTAAATGCGACGATACCTGTCCAAAGCCGTACTT 600
Db 181 GlnGlnAsnIuYrIleGlyIuYrValLeuAsnAlaSerAsnThrCyPheProIuYrAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGGTGCTGTGATAGACAGCTTTGTGAAATTTTGG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCyGlnIuYrIleLeu 220

QY 661 CGGTGTGGACAATCAGATGATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTT 720
Db 221 ProTrpLeuAspIuYrSerValAspPheIleGlnIuValAlaValAsnAsnIuYrVal 240
QY 721 CTAGTCACTGTTTACCTGGATCTCCGCTCCGACCATCTGCTATGCTTACATCATG 780
Db 241 LeuValHisCyLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIuYrIleMet 260
QY 781 AAGAGATGAGACATGCTTTTACATGAACTTACAGATTTGTGAAAGAAAAAGCTTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnIuAlaIuYrArgPheValIuYrGlnIuYrAspProThr 280
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGACTATGAGAGAAAGATTAAGAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlnIuLeuAsnProIuYrIuYrIleYrAsn 300
QY 901 CAGACTGAGCATCAGAGCCAAAGACAACTCAAGCTGTGACCTGAGAGCCCAAT 960
Db 301 GlnThrGlyAlaSerGlyProIuYrSerIuYrLeuIuYrLeuIuYrIleGlnIuYrProAsn 320
QY 961 GAACTGTCCCTGCTGTCTTCAGAGGTTGACAGAAAGGACGCCCTCACTCACTCC 1020
Db 321 GluProValProAlaValSerGlnIuYrGlnIuYrSerGlnIuYrProIuYrSerProPro 340
QY 1021 TGTGCGAGCTGTCTACCTCAAGAGGACAGACAAAGCCGCTGATCCGACAGCTG 1080
Db 341 CyAlaAspSerAlaThrSerGlnAlaIleGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCGAGGCTCCCGACGCTGACAGCCGCTGTTTGAAGACAGCCGCTGTATACAGCGCTC 1140
Db 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGACCTGTCTCCGACAGAGCTGAAAGACAAATAGCTCAAGCTTCTTTC 1200
Db 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnIuAspSerAsnIuYrLeuYrAspSerPhe 400
QY 1201 TCTCTGATATCAATACATGTTTCAATTCAGCAGAGTGCACATCTTACATAGCTTC 1260
Db 401 SerLeuAspIleYrSerValSerIuYrSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
QY 1261 TCTCATCAGAGATCTTGGATCTCAAACTTCCATCTGATGAGAGCCCAAC 1320
Db 421 SerSerSerGlnAspAlaLeuGlnIuYrIuYrIuYrProSerThrIleuAspGlyThrAsn 440
QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATTCGAGAGACCTCCGAAACAGTCT 1380
Db 441 LysLeuCyGlnPheSerProValGlnIuLeuSerGlnIuYrProGlnIuYrThrSerPro 460
QY 1381 GATTAAGAGAAAGCCAGCATCCCAAGAGCTGCACACCGCCAGGCTTCAGACAGCAG 1440
Db 461 AspIuYrGlnIuAlaSerIleProIuYrIuYrLeuGlnIuYrAlaArgProSerAspSerGln 480
QY 1441 AGCAAGCATGTCATTCGCTCAGAACTGAGAGACAGAGAGCAATTCACACCACTTCTTTCG 1500
Db 481 SerIuYrGlnIuHisSerValArgIuThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
QY 1501 TCTCACTGCATCAAGTGGAGCTGAGAGACATTCACACCACTTCTTTCG 1560
Db 501 SerProLeuHisArgSerGlySerValGlnAspAsnIuYrHisPhePheLeuPheGly 520
QY 1561 CTTCACACAGCAGAGACAGCATCCAGAGTGTGCTGGCTGCTTAAAGGCTGGCAG 1620
Db 521 LeuSerThrSerGlnIuHisLeuIuYrIuYrSerAlaGlyLeuIuYrLeuYrGlyIuYrPheHis 540
QY 1621 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTTACACAGAGCTGTATTTTGGC 1680
Db 541 SerAspIleLeuAlaProGlnIuYrIuYrIuYrProSerLeuThrSerSerIuYrIuYrPheAla 560
QY 1681 ACAGAGTCTTCACTTACTGCTGACCATTCAGAGAGCAGTGCAGTACTCT 1740
Db 561 ThrGlnSerSerHisPheIuYrSerAlaSerAlaIleIuYrGlyIuYrSerAlaSerIuYrSer 580


```

QY 661 CCGTGGTGGACAAATCATAGATTTCATTGAGAAAGCAAAAGCCTCAATGGATGTT 720
    |||
Db 221 ProTribLeuApblysservAlaephneIIeGIuYalAlaYsaIaseraHngIYCsVal 240
QY 721 CTAGTCACTGTTTAACTGGGAATCTCCCGCTCCGCCACATCGCTATCGCTCATCATG 780
    |||
Db 241 LeuValHieCyLeuAlaGIYIleSerArgSerAlaThrIleAlaIleAlaIleAlaIleMet 260
QY 781 AAGAGATGAGACATGCTTTCATGATGAAAGCTTACAGATTGTGAAAGAAAAAGACTACT 840
    |||
Db 261 LysArgMetAspMetSerLeuApblyuAlaYArArgPheValIysGIuYsaGProThr 280
QY 841 ATATCTCCAAATCTCAATTTCTGGGCAACTCGTGCATATGAGAAAGAAATTAAGAAC 900
    |||
Db 281 IleSerProAsnPheAsnPheLeuGIYGIuLeuLeuApblyrGIuYsIleYsAsn 300
QY 901 CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGAGAAAGCAAT 960
    |||
Db 301 GIuThrGIYAlaSerGIYProLYsSerLYeuleuLYeuleuHieIeugIuYsProAsn 320
QY 961 GAACCTGCTCCCTGCTGCTCAGAGGGTGAACAGAAAAAGCAGACCCCTCATCCACCC 1020
    |||
Db 321 GIuProValProAlaValSerGIuGIYGIuLYsSerGIuThrProLeuSerProPro 340
QY 1021 TGTGCGGACTGCTGCTACCTCAGAGGAGCAGACAAAGGCCGCTGCATCCCGCAGGCTG 1080
    |||
Db 341 CyAlaIAspSerAlaThrSerGIuAlaAlaGIYGIuThrProValIAspProAlaSerVal 360
QY 1081 CCCAGCGTGGCCAGCGCTGACAGCGCTGCTGTTAGAGAGACGCCGCTGTATACAGCGCTC 1140
    |||
Db 361 ProSerValProSerValGIuProSerLeuLeuGIuAspSerProLeuValGIuAlaLeu 380
QY 1141 AGTGGGCTGACCTGCTCCGAGACAGGCTGGAAGACGACATTAAGCTCAAGCTTCTTC 1200
    |||
Db 381 SerGIYLeuHieIeuleuSerAlaAspArgLeuGIuAspSerAsnLYsIeuleuYsaGSerPhe 400
QY 1201 TCTCTGATATCAATCATGTTTATATTCAGCAGACGTGCACATCCTTACATGGCTTC 1260
    |||
Db 401 SerLeuApblyIeYsservAlaSerTYrSerAlaSerMetAlaIAserLeuHieGIYpHe 420
QY 1261 TCCTCATCAGAAAGATGCTTGAATATCTAACAACCTTCCATCTGTGATGGAGCAAC 1320
    |||
Db 421 SerSerSerGIuApblyAlaLeuGIuTYrTYrLYeProSerThrIeuleuApblyThrAsn 440
QY 1321 AAGCTATGCCAGTTCCTCCTGTTCAAGAACTATCGAGAGACACTCCGCAACCACTGCT 1380
    |||
Db 441 LYsLeuCyseGIuPheSerProValGIuGIuLeuSerGIuGIuThrProGIuThrSerPro 460
QY 1381 GATTAAGAGAGAGACGACATCCCAAGAAAGCTGACAGCCGCGCTTACAGACGACAG 1440
    |||
Db 461 AspLYsGIuGIuAlaSerIleProLYsLYsLeuGIuThrAlaArgProSerAspSerGIu 480
QY 1441 AGCAAGCGATTGCTCGGTCAAGAACGACAGACAGTGCACCGCCGAGAGTCCCTTTTA 1500
    |||
Db 481 SerLYsArgLeuHieSerValaArgThrSerSerSerGIYThrAlaGIuArgSerLeuLeu 500
QY 1501 TCTCCATCGCATGAAGTGGAGCGGTGAGAGCAATTAACAACGACTTCTTTTCGGC 1560
    |||
Db 501 SerProLeuHieArgSerGIYSerValGIuAspAsnTYrIleThrSerPheLeuPheGIY 520
QY 1561 CTTTCACACGACGAGACGACCTCAGAGAGTCTGCGGCTGCGGCTTAAAGGCTGGCAC 1620
    |||
Db 521 LeuSerThrSerGIuGIuHieIeuleuThrLYsSerAlaGIYLeuGIYLeuLYsGIYTPHAs 540
QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACGACGAGCTGTATTTGGC 1680
    |||
Db 541 SerAspIleLeuAlaProGIuThrSerThrProSerLeuThrSerSerTYrPheAla 560
QY 1681 ACAGAGTCTCTCAACTTCTACTGCTCAGCCATCTACGAGAGGACGAGTCACTTCTCT 1740
    |||
Db 561 ThrGIuSerSerIlePheTYrSerAlaSerAlaIleTYrGIYsIeuleuAlaSerTYrSer 580
QY 1741 GCCTACAGCTGACGACGAGCTGCCACTGTGGGAGAGCAAGTCAATTCTGTGGCAGGCG 1800

```

```

Db 581 AlaTYrSerCySerGIuLeuProThrCYsGIYAspGIuValYrSerValaArgArg 600
QY 1801 CAGAACCCAAAGTACAGAGCTGACTCGCGCGGAGCTGGCATTAAGAGACCCCTTTGAA 1860
    |||
Db 601 GIuLYsProSerAspArgAlaAspSerArgArgSerTYrHieGIuGIuSerProPheGIu 620
QY 1861 AAGCACTTTAAACGAGAGAGCTGCCAAATGGAATTTGGAGAGACATCATGTCAAGAAC 1920
    |||
Db 621 LYsGIuPheLYsArgArgSerCYsGIuMetGIuPheGIYGIuSerIleMetSerGIuAsn 640
QY 1921 AGGTCAAGGAGAGAGCTGGGGAAGTGGGACAGTCACTTACCTTTTCGGGACAGAGAA 1980
    |||
Db 641 ArgSerArgGIuGIuLeuGIYsValaGIYsSerGIuSerSerPheSerGIYsSerMetGIu 660
QY 1981 ATCATTAAGCTCTCC 1995
    |||
Db 661 IleIleGIuValSer 665
    |||

RESULT 5
AAU79156
ID AAU79156 standard; protein; 665 AA.
XX
AC AAU79156;
XX
DT 02-JUL-2002 (first entry)
XX
DE Human dual-specificity phosphatase-3 (DSP-16) protein.
XX
KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; chromosome 12p.
XX
OS Homo sapiens.
XX
FH
FT Key Location/Qualifiers
FT Domain 242..251 /label= Active_site_domain
PN MO200226997-A2.
XX
PD 04-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US030124.
XX
PR 26-SEP-2000; 2000US-0235487P.
XX
PA (CEPT-) CEPTYR INC.
XX
PI Luche RM, Wei B;
XX
DR N-PSDB; ABK47596.
XX
DR MPI; 2002-315802/35.
XX
PT New DSP-16 polypeptide, useful for identifying modulators of its
PT activity, which can be used in the treatment of disorders such as
PT Duchenne muscular dystrophy, or cancer.
XX
PS Claim 1; Fig 2; 87pp; English.
XX
CC The present invention relates to a new polypeptide, DSP-16, having a 665
CC amino acid sequence, given in the specification, or a variant having at
CC least 50 % identical residues, which retains the ability to
CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
CC invention can be used for identifying agents which modulate DSP-16
CC activity, for modulation of a proliferative response in a cell, survival
CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne

```


QY	1921	AGGTCACCGGAAAGAGCTGGGGAAAGTGGGAGTCAGTCTACGTTTTCGGGAGCATGGAA	1980
Db	641	Argseraiaagglugluuueugllylvaldylserginserserpheserdcylsermetcu	660
QY	1981	ATCATTTGAGGTCTCC	1995
Db	661	lllellguValser	665
RESULT 6			
ID	AAU09946	standard; protein; 665 AA.	
XX	AAU09946;		
XX	18-JUN-2002	(first entry)	
DE	Protein sequence of human (dual specificity phosphatase) DUSP-10.		
XX	Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;		
KW	neuronal degeneration syndrome; Alzheimer's disease; depression;		
KW	schizophrenia; asthma; immune disorder; inflammatory process; arthritis;		
KW	osteoporosis; diabetes.		
OS	Homo sapiens.		
XX			
PN	WO2001.77340-A1.		
XX	18-OCT-2001.		
XX	06-APR-2001; 2001WO-EP003966.		
XX	10-APR-2000; 2000EP-00107143.		
XX	(MERE) MERCK PATENT GMBH.		
PA			
XX			
XX	Duecker K;		
XX	WPI; 2002-010917/01.		
DR	N-PSDB; AAS15768.		
XX			
PT	Novel dual specificity phosphatase polypeptides useful for treating		
PT	cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's		
PT	disease, depression, schizophrenia, asthma and immune disorders.		
XX			
PS	Claim 2; Page 37-39; 43pp; English.		
XX			
XX	The present invention relates to a new isolated dual specificity		
CC	phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid		
CC	sequence that is fully defined in the specification. The invention also		
CC	provides a sequence encoded by a 3059 nucleotide sequence fully defined		
CC	in the specification, and a sequence having at least 95 % identity to the		
CC	polypeptide, or fragments or variants of DUSP-10. The invention is useful		
CC	for treating cancer e.g. leukemia, colon carcinoma, lung cancer,		
CC	prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,		
CC	stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,		
CC	schizophrenia, cardiac myopathies, asthma, immune disorders,		
CC	inflammatory processes e.g. arthritis, bowel disease, type I diabetes,		
CC	osteoporosis, diabetes and diabetes associated diseases. The molecules of		
CC	the invention are also useful as vaccines for inducing immunological		
CC	response in a mammal, in disease diagnosis and in assays for screening		
CC	agonistic or antagonistic compounds. Other uses of the invention include		
CC	identifying membrane bound or soluble receptors, as a diagnostic reagent,		
CC	in chromosome localisation studies, and as a valuable tool in tissue		
CC	expression studies. The present sequence represents the dual specificity		
CC	phosphatase, DUSP-10, protein of the invention		
XX			
XX	Sequence 665 AA;		
XX			

Alignment Scores:	3.16e-287	Length:	6655
Pred. No.:		Matches:	6653
Score:	3406.00	Conservative:	1
Percent Similarity:	99.85%		

Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	94.22%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-108_COPY_538_2532 (1-1995) X AAU09946 (1-665)

```

Mismatches: 1
Indels:      0
Gaps:        0

```

QY	ATGCCCATGAGATGATTTGGAACTCAATTTGTACTGAGAGTTGGTGGCTCTGGAGA	60
Db	1 MetAlaHisGluMetCileGlyThrGlnIleValThrGluArgLeuValAlaIleuLeuGlu	20
QY	61 AGTGGAA CGGAAAAAGTGCCTGCTAATTGATAGCGGCGCATTTGTGGAATACATACATCC	120
Db	21 SerGlyThrGluArgValIleuLeuIleAspSerArgProPheValGluArgAsnThrSer	40
QY	121 CACATTTTGGAAAGCCATTATATCACTGCTCCAAAGCTTATGAAAGCGAAGTTTGCACAG	180
Db	41 HisIleuGluAlaIleAsnIleAsnCySerIysLeuMetLysArgAlaGluGln	60
QY	181 GACAAAGTGTTAATTACAGAGCTCATCGAGATTCACAGAAACCTAAAGTTGACATTGAT	240
Db	61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysIleValAspIleAsp	80
QY	241 TGCAGTCAGAAAGTTGATTTAGATTACGATCAAAAGCTCCAGATGTTGCCCTCTCTCTTCA	300
Db	81 CysSerGlnLysValValIleValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer	100
QY	301 GACTGTTTCTCACTGATTACTTCGGGTAACTGGAGAAAGCTTCAACTCTGTTCACTCG	360
Db	101 AspCysPheLeuThrValIleuLeuGlyLysLeuGluLysSerPheAsnSerValHisIleu	120
QY	361 CTTTCAGAGTGGGTTTGGCTGAGTTTCTCGTGTGTTTCCCTGGCCTCTGTGAAGAAATCC	420
Db	121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCyGluGlyLysSer	140
QY	421 ACTCTAGTCCCTACCTGCATTTCTCAGCCTTGCTTACCTGTTGCCAATTTGGGCCAAC	480
Db	141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
QY	481 CGAATTTCTTCCCAATCTTTATCTTGGGTCGACAGAGATGTCCTCAACAGAGCTGATA	540
Db	161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuMet	180
QY	541 CAGCAGATGGAGATTTGGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT	600
Db	181 GlnGlnAsnGlyIleGlyTyrValIleuAsnAlaSerAsnThrCysProLysProAspPhe	200
QY	601 ATCCCGAGTCTCATTTCTCGCGTGGCTGTGAATGACAGCTTTTGTGAGAAATTTTG	660
Db	201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu	220
QY	661 CCGTGTTTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGGCTCCAAATGAGATGTTT	720
Db	221 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysValAspAsnGlyCysVal	240
QY	721 CTAATGCACTGTTTAGCTGGAGATTCGCCGCTCCGCAACATCGCTATCGCTTACATCATG	780
Db	241 LeuValHisCysLeuAlaGlyIleSerHisSerHisAlaThrIleAlaIleAlaTyrIleMet	260
QY	781 AAGAGATGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTACT	840
Db	261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr	280
QY	841 ATATCTCCAAACTCATATTTTCTGAGGCAACTCTGTGACATTAAGAGAAATTAAGAAC	900
Db	281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLysAsn	300
QY	901 CAGACTGAGCATGAGGCGCAAAAGCAAACTCAAGTCTGCACTTGGAGAGCCCAAT	960
Db	301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGluLysProAsn	320
QY	961 GAACCTGTCCTGCTGCTCAGAGGGGTGGACAGAAAAAGCAGAGCGCCCTCAGTCCACCC	1020
Db	321 GluProValProAlaValSerGluGlyGlyGlnLysSerGlnThrProLeuSerProPro	340

QY	1021	TGTCGCCACTGTGGTCACTTGACAGCGCAGACGAACAAGCCGCCTGATCCCCGACCGTTG	1087
Db	341	CysAlaAspSerSeraIatTrSerGIuaIaaIagIyGlnAArgProValHISPrOlaIsarVal	360
QY	1091	CCAGCGCTGGCCACCGCGCAGCGCGTGGCTTTAAAGAACACGCCGCTGGTAACAGCGCTC	1140
Db	361	ProSerValProSerValGlnProSerLeuLeuGlnAspSerProleuValGlnAlaLeu	380
QY	1141	AGTGGCGTGCACCTGTGCCCAACACAGCGCTGGAGAAGCAATAAGCTCAAGGGCTTCCTC	1200
Db	381	SerCIlyeunHISleuSerAlaaspArgLeuGluAspSerHnlySpSerHnlySarSerPhe	400
QY	1201	TCTCTGGATATCAAATCAAGTTTCATATTACGCCGACGATGGCACATCTTTACATGGCTTC	1260
Db	401	SerLeuAspIlElySerValSerTySeraIasrMetAlalaserLeuHISglYphe	420
QY	1261	TCTCATTCAGAGAGATGCTTTGGAAATCTACAAACTTCCATCACTCTGATGGAGACCAAC	1320
Db	421	SerSerSerGluAspAlalaLeuGluTyTyLysProSerThrThIleuAspGlyThnAsn	440
QY	1321	AAGCTATGCGAATTCTCCCCTGTTCAGAACTATCGAGACGACATCTCCGAAACCAGTCTC	1380
Db	441	LysLeuCysGlnPheSerProvalGlnGluLeuSerGlnGlnInThrProGluInThrSerPro	460
QY	1381	GATAAGAGAGAAAGCCACATCCCAAGAACTGTCAGACCGCGACGCTTCAGACAGCGAC	1440
Db	461	AspLySGluGluAlaSerIlleProLySLysLeuGlnThrAlaArgProSerAspSerGln	480
QY	1441	AGCAAGCGATTGCATTGGATCAGAACACAGACACATGGACACCGGCCAGAGTCCCTTTTA	1500
Db	481	SerLysArGLeuHISserValArgThrSerSerSerGlyThnAlaGlnArgSerLeuLeu	500
QY	1501	TCTCCACTGCATCGAAGTGGAGCGCTGGAGAGCAATTACCAACCAAGCTCTCTTTTCGCG	1560
Db	501	SerProLeuHISargSerGlySerValGlnUspsanLyThISthrSerPheLeuHecly	520
QY	1561	CTTTCCACCGCCAGACGACCTTCACGAAGTCTGCTGGCGCTTAAAGGCTGGGAC	1620
Db	521	LeuSerThrSerGlnInHISleuThrLysSerAladilyeudlyeunylslyTriPhis	540
QY	1621	TGGGATATCTTGAGCCGCCACAGCCCTTACCCCTTCCCTGACACGACGTGGATTTTTGCC	1680
Db	541	SerAspIlleAlaValProGlnThrSerThrProSerLeuThrSerSerTriPhyPhala	560
QY	1681	ACGAGATCTCTCACTTTTACTCTTGCTCAGCCATTCAGAGAGCAGAGTGAATTACTCT	1740
Db	561	ThnglUSeSerHisPheTySerAlaSerAlaIleTyGlyGlySerAlaSerTySer	580
QY	1741	GCCTACAGCTGCAGCGACGTGCCCACTMGGGAGAGCAAGTCTATTCTGTGCGACAGCGG	1800
Db	581	AlaTySerSerCySerGlnLeuProThrCySGlyAspGlnValTySerValArgArgArg	600
QY	1801	CAGAAGCCAATGACACAGCTGACTGCGGCGGAGACTGGCATGAAGAGACCCCTTTGAA	1860
Db	601	GlnlySProSerAspAlaAspSerArgAspSerTriphISglnGluusertPropHeGlu	620
QY	1861	AAGCATTTAAACGACAGAGCTGCCAAATGGAATTGGAGAGAGCATCATGTCAAGAAC	1920
Db	621	LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIlleMetSerGluSn	640
QY	1921	AGGTCACGGAAAGGTGGGGGAAAGTGGGACGTCAAGTCTTAGCTTTTCGGGCGACATGGAA	1980
Db	641	ArgSerThrGlnGluInuGlnLysValGlySerGlnSerSerPheSerCIlyserMetGlu	660
QY	1981	ATCATTTGAGGTCTCC	1995
Db	661	IleileGluValSer	665

AC	AAU5789;	
XX		
DT	08-MAY-2002 (first entry)	
XX		
DE	Human protein phosphatase 7 (PP7) protein sequence.	
XX		
KW	Human; protein phosphatase 7; PP7; immune system disorder; AIDS;	
KW	acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma	
KW	Crohn's disease; neurological disorder; epilepsy; Huntington's disease	
KW	dementia; Parkinson's disease; developmental disorder; Down's syndrome	
KW	cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymph	
XX	melanoma; myeloma sarcoma.	
OS	Homo sapiens.	
XX		
Key	Location/Qualifiers	
FH	Domain	11..131
FT	/label="Rhodanese_like_domain	15..170
FT	Region	/note="Dual specificity protein phosphatase"
FT	Region	85..298
FT	Domain	/note="VH1-type dual specificity protein phosphatase"
FT		158..297
FT	/label="Catalytic domain DSPc	220..281
FT	/note="Dual specificity phosphatase"	
FT	Region	/note="Tyrosine specific protein phosphatases active
FT	Region	site tyr phosphatase.pif"
FT	Region	242..254
XX	/note="Tyrosine phosphatase"	242..252
PN	/note="Tyrosine specific protein phosphatases"	
XX		
PD	MO200210363-A2.	
XX		
07-FEB-2002.		
XX		
26-JUL-2001; 2001WO-US023716.		
XX		
28-JUL-2000; 2000US-0221679P.		
PR	03-AUG-2000; 2000US-0223722P.	
PR	10-AUG-2000; 2000US-0224309P.	
PR	18-AUG-2000; 2000US-0226728P.	
PR	30-AUG-2000; 2000US-0229254P.	
PR	08-SEP-2000; 2000US-0231366P.	
XX		
XX	(INCY-) INCYTE GENOMICS INC.	
PI	Tang YF, Elliott VS, Rankumar J, Yao MG, Burford N, Wang YE;	
PI	Stewart EA, Gandhi AR, Patterson C, Lee EA, Hatalla AJA, Lu DM;	
PI	Trilouley CM, Griffin JA, Baugh MR, Yue H, Warren BA, Nguyen DB;	
XX	Walia NK, Kearney L;	
XX		
DR	MPI; 2002-188735/24.	
XX		
DR	N-PDSB; ABK14474.	
PT	New protein phosphatases, useful for diagnosing, treating or preventing	
PT	immune system disorders (e.g. Crohn's disease), neurological disorders	
PT	(e.g. Parkinson's disease), or cell proliferative disorders (e.g.	
PT	cancers).	
PS	Claim 1; Page 107-108; 117pp; English.	
XX		
CC	The present invention relates to a new polypeptide, a naturally occurring	
CC	amino acid sequence at least 95 % identical to it, a biologically active	
CC	fragment of it or an immunogenic fragment of it, the polypeptides,	
CC	polynucleotides, agonists and antagonists are useful for diagnosing,	
CC	treating or preventing disorders associated with aberrant expression of	
CC	protein phosphatases (PP), particularly immune system disorders e.g	
CC	acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,	
CC	asthma or Crohn's disease, neurological disorders e.g. epilepsy,	
CC	Huntington's disease, dementia or Parkinson's disease, developmental	
CC	disorders e.g. Down's syndrome, or cell proliferative disorders e.g.	

Db 641 ArgSerArgGluGluLeuGlyValGlySerGlnSerSerPheSerGlySerMetGlu 660
Qy 1991 ATCATTTGAGCTCTCC 1995
Db 661 IleIleGluValSer 665
RESULT 8
ID ABB97946 standard; protein: 665 AA.
XX ABB97946;
AC ABB97946;
XX 06-SEP-2002 (first entry)
DB Human protein sequence #13.
XX Human protein sequence #13.
XX Human; brain; tonsil; hippocampus; foetal brain; diagnosis.
OS Homo sapiens.
XX M0200252005-A1.
PN 04-JUL-2002.
XX 20-DEC-2001, 2001MO-JP011217.
PF 22-DEC-2000, 2000JP-00389742.
PR (KAZU-) KAZUSA DNA RES INST FOUND.
XX (CELE-) CELESTAR LEXICO-SCI LTD.
PA Ohara O, Nagase T, Nakajima D;
PI WPI: 2002-500762/53.
XX N-PSDB; ABB83966.
DR Genes and their expression products cloned from human cDNA libraries for
XX treatment and diagnosis of diseases associated with their expression.
PT Claim 1(a); Page 112-116; 238pp; Japanese.
XX
XX The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies,
CC oligonucleotide probes and DNA chips for diagnosis and identification of
CC drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABB97934-ABB97964 represent
CC human proteins of the invention
XX
SQ Sequence 665 AA;
Alignment Scores:
Pred. No.: 3,166-287 Length: 665
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 94.22% Indels: 0
DB: 5 Gaps: 0
US-10-029-345a-108_copy_538_2532 (1-1995) x ABB97946 (1-665)
Qy 1 ATGGCCGATGAGGATTTGGAACCTCAATTTGTTACTGAGAGGTTGGTGGCTGCTGAGAA 60
Db 1 MetAlaHisGluMetCileGlyThrGlnIleValThrGluArgLeuValAlaIleuLeuGlu 20
Qy 61 AGTGAACGAGAAAAGTGTCTTAATTGATAGCGGCGCATTTGTGGAATACATACATCCTCC 120
Db 21 SerGlyThrGlyLeuValIleuLeuIleAspSerArgProPheValGluIleuThrSer 40

Qy 121 CACATTTTGAAGCCATTATATCACTGCTCCAAAGCTTATGAGCCGAAGTTGCAACAG 180
Db 41 HisIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 60
Qy 181 GACAAAGCTTAAATTAACAGAGCTCATCCAGCATTCAGCAAAACATAAGTTGACATTGAT 240
Db 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaIleHisValAspIleAsp 80
Qy 241 TGCAGTCAGAGGTTGTAGTTTACATCAAGAGCTCCGAGATGTCCTCTCTCTTCA 300
Db 81 CysSerGlnLysValIleValIleValIleValIleValIleValIleValIleValIle 100
Qy 301 GACGTTTTCACCTGATCTTCTGGGTTAACTGAGAAAGCTTCAACTGTGTCACCTG 360
Db 101 AspCysPheLeuIleThrValIleLeuGlyLysLeuGluLysSerPheAsnSerAlaHisLeu 120
Qy 361 CTTCGAGGTGGGTTTGTGAGTTCTCTCGTTGTTTCCGCGCTGCTGTAAGAAATTC 420
Db 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluIleLysSer 140
Qy 421 ACTGATGCTCCATCCTGCAATTTCTCAGCCTTGCTTACCTGTGCAAGATTGGGCAAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Qy 481 CGAATTTTCCCAATCTTATCTTGCTGCTGCGAGAGATGTCCTCAACAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuIleGlyLysGlnIleGlyAspValIleuAsnLysGluLeuMet 180
Qy 541 CAGAGAAATGAGATGCTGATGTTATGTTAAATGCAAGCTTACCTGCAAGAGCTGACTTT 600
Db 181 GlnGlnAsnGlyIleGlyLysValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
Qy 601 ATCCCGAGTCTCATTTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 201 IleProGlnSerHisPheLeuAlaGlyValProValAsnAspSerPheCysGluLysIleLeu 220
Qy 661 CCGTGGTTGGAACAATTCAGTGAATTTTCAATGAGAAACCAAAAGCTTCAATGATGTTG 720
Db 221 ProThrPheAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
Qy 721 CTATGTCACGTGTTAGCTGGAGTCTCCGCTCCGCAACATGCTTACCTTACATCAGT 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIle 260
Qy 781 AAGAGAGTGAACATCTCTTATGATGAAGCTTACAGATTTGTGTAAGAAAGAAAGCTTACT 840
Db 261 LysArgMetAspMetSerLeuAspGluAlaIleValIleValIleValIleValIleVal 280
Qy 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGCACTATGAGAAAGAAAGTTTAAGAAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlnIleuLeuAspIleGluLysLysIleLysAsn 300
Qy 901 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGCACTGAGAAAGCAAT 960
Db 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisIleGluLysProAsn 320
Qy 961 GAACCTGCTCCCTGCTGCTGCAAGGCTGAGCAAGAAAGCAAGCAAGCAAGCAAGCAAG 1020
Db 321 GluProValProAlaValSerGluGlyGlnLysSerGlnThrProLeuSerProPro 340
Qy 1021 TGTGCGGATCTGCTTCACTCAGAGCAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1080
Db 341 CysAlaAspSerAlaHisSerGlnAlaIleGlyGlnArgProValHisProAlaSerVal 360
Qy 1081 CCGAGGTCCTCAGAGGTCAGAGCTGCTGCTGTTAGAGCAAGCCGCTGCTGCAAGGCTC 1140
Db 361 ProSerValProSerValGlnProSerIleuLeuGlnAspSerProLeuValGlnAlaLeu 380
Qy 1141 AGTGGGCTGCACTCTCTCCGAGCAAGGCTGGAAGCAAGCAATTAAGCTCAAGCTTCTTC 1200
Db 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 400
Qy 1201 TCTTGATATCAATCATGTTTCATATTCAGCCAGCATGCGACAGATCTTACATGCTTC 1260

Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 94.22%
DB: 5

Conservative: 1
Mismatches: 1
Indels: 0
Gaps: 0

US-10-029-345a-108_COPY_538_2532 (1-1995) x AAU97929 (1-665)

```
QY 1 ATGGCCCATGAGATGAGTGAACCAATTGTTACTAGAGGTTGGTCTGCTGGAA 60
DB 1 MetalaHisglumetIleglyThrglnIlevalThrglnArgleuValAlaleuEnglu 20
QY 61 AGTGAACGGAAAAAGCTGCTAATTGATGACCGGCAATTTGTGGAATAATCATCC 120
DB 21 SerGlyThrglnValIleuLeuIleAspSerArgProPheValGlnIyrzsmthsr 40
QY 121 CACATTTTGAAGCATTATATCACTGCTCAACCTTATAGAGCAAGTTGGACAG 180
DB 41 HisIleuEngluAlaleuIleamCysSerLysleuMetLysArgArgleuGln 60
QY 181 GACAAAGTTTATTTACAGACTCATCCAGCATTCAGCAAAACATAGATTGATGAT 240
DB 61 AspLysValIleuIleThrglnleuIleGlnHisSerAlaLysIleLysValAsp 80
QY 241 TGCAGTCAGAAAGTTTGTAGTTTACAGTCAAGCTCCCAAGTGTGCTCTCTTCA 300
DB 81 CysSerGlnLysValAlaValIyrzAspGlnSerSerGlnAspValAlaSerLeuSer 100
QY 301 GACTGTTTCTGACGTACTTCTGGGTAACCTGAGAAAGACTCACTGTTTCACTG 360
DB 101 AspCysPheLeuThValIleuLeuLysLysLysSerPheamSerValHisLeu 120
QY 361 CTTCAGAGTGGTGTGCTGAGTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProLysLeuCysGlnGlySer 140
QY 421 ACTTAGTCCCTACCTGATTTCTGACCTTGTCTTACCTGTTGCCAAATGGGCAACC 480
DB 141 ThriLeuValProthnCysIleSerGlnProCysLeuProValAlaAsnIleGlyProth 160
QY 481 CGAATTTCTCCCAATCTTTATCTTGCTGCTGACGAGATGCTCTCAACAGAGCTGTA 540
DB 161 ArgIleuProLeuLeuIyrleuGlyCysGlnArgAspValLeuAsnIleGlyProth 180
QY 541 CAGAGAAATGGGATGTTATGTTTAAATGCCAGTATACCTGTCCAAGGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyIyrValIleuAsnAlaSerAsnThrCysProLysPheAsp 200
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGAGATGACAGCTTTTGTGAAGAAATTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleu 220
QY 661 CCGGTTTGGACAATACTAGATTTTATTGAGAAAGCAAAAGCTCCATGATGATGTT 720
DB 221 ProThrPheAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTATGCTACTGTTTACCTGGATCTCCGCTCCGCAACCATCGCTATCGCTTACATCAG 780
DB 241 LeuValHisCysLeuAlaGlyIyrleSerArgSerAlaThrIleAlaIleAlaIyrIleMet 260
QY 781 AAGAGATGAGCATGCTTTAGTAGAGCTTACGATTTGTGAAGAAAAAGCACTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaIyrzArgPheValLysGlnLysArgProth 280
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCATAGAGAAAGATTTAAGAAC 900
DB 281 IleSerProLeuPheAsnPheLeuGlyGlnLeuLeuAspIyrGlnLysLysIleLysAsn 300
QY 901 CAGACTGAGCATGAGGCGCAAGAGCAAACTCAAGCTGCTGACCTGGAGAGCAAT 960
DB 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisIleuGlnLysSerAsn 320
QY 961 GAACCTGCTCTGCTGCTCAGAGGTTGACAGAAAGGAGAACCCCTCATGTCACCC 1020
```

```
DB 321 GlnProValProAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerProPro 340
QY 1021 TGTGCCAGCTCTGTACTTACCTCAGAGGCAAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
DB 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGCTGCCAGGCTGACGCGTCTGTTTATAGAGCAAGCCCGCTGTGACAGCGCTC 1140
DB 361 ProSerValProSerValGlnProSerLeuLeuLysAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGCACTCTGTCGCGAGACAGGCTGGAGACAGCAATAAGCTCAAGCTTCTTC 1200
DB 381 SerGlyLeuHisIleuSerAlaAspArgleuGlnAspSerAsnLysLeuLysArgSerPhe 400
QY 1201 TCTTGATATCAATCAATGATTTCATATTACGACAGCATGGAGCATCTTACATGGCTTC 1260
DB 401 SerLeuAspIleLysSerValSerIyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
QY 1261 TCTCATCAGAAAGATGCTTTGGAATCTACAAACTTCCACTACTCTGATGGAGCAAC 1320
DB 421 SerSerSerGlnAspAlaLeuGlnIyrIyrLysProSerThrThrLeuAspGlyThrAsn 440
QY 1321 AAGTTATGCCAGTTTCTCCCTGTTTCAAGAACTATGCGAGCAAGACTCCGGAACAGTCTT 1380
DB 441 LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnGlnIyrProGlnThrSerPro 460
QY 1381 GATTAAGAGAGAACCCAGCATCCCAAGAAAGCTGCAAGCCGCGAGCTTCCAGACGACAG 1440
DB 461 AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
QY 1441 AGCAACGATTTGCTTGGTCAACACAGACAGAGTGGACCGGCGAGAGTCCCTTTTA 1500
DB 481 SerLysArgLeuHisSerValAlaArgThrSerSerSerGlyThrAlaGlnAspSerLeu 500
QY 1501 TCTCCACTCATCAAGAGTGGAGCGTGGAGAGCAATTAACAACACAGCTTCTTTTCCG 1560
DB 501 SerProLeuHisArgSerGlySerValGlnAspAsnIyrzHisThrSerPheLeuPheGly 520
QY 1561 CTTTTCCACGACGACAGACACTTCAAGAGCTGCTGCTGGGCGCTTAAAGGCTGGCAC 1620
DB 521 LeuSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuGlyLeuLysGlyIyrzHis 540
QY 1621 TCGATATCTTGGCCCCCGACGACTTACCCCTTCCCTGACACAGCAGCTGTATTTTGC 1680
DB 541 SerAspIleLeuAlaProGlnIyrThrProSerLeuThrSerSerIyrzPheAla 560
QY 1681 ACAAGATCTTCAACTTCTACTGCTCAGCATCTACGAGGACAGTCCAGTACT 1740
DB 561 ThrGlnSerSerHisPheIyrSerAlaSerAlaIleIyrGlyGlySerAlaSerIyrSer 580
QY 1741 GCTTACAGCTGCAAGCAGCTGCCCACTTGGCGAGAACCAAGTATTCTGTGCGACGCG 1800
DB 581 AlaIyrSerCysSerGlnLeuProthnCysGlyAspGlnValIyrSerValArgArg 600
QY 1801 CAGAAGCAAGTACAGAGCTGACTCGCGCGAGACTGGCATGAAGAGAGCCCTTTGAA 1860
DB 601 GlnIyrProSerAspArgAlaAspSerArgArgSerIyrzHisGlnLysSerProPheGln 620
QY 1861 AAGCAGTTTAAACGCAAGAGCTGCCAAATTTGAGAGAGACATCATGTACAGAAC 1920
DB 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGGTCAACGGGAAGACTGGGGAAGTGGGAGTCAAGTCAAGTCTTTCGGGACATGGA 1980
DB 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATTTGAGTCTCC 1995
DB 661 IleIleGlnValSer 665
```

RESULT 10
ABB97291
ID ABB97291 standard; protein; 665 AA.

XX ABB97291;
AC
XX
XX 27-JUN-2002 (first entry)
XX
XX
XX Novel human protein SEQ ID NO: 559.
XX
XX Human, anti-nausea; vulnerary; anti-inflammatory; immunomodulator;
XX anti-infective; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
XX neuroprotective; anti-Parkinsonian; protein therapy; EST;
XX expressed sequence tag.
XX
XX Homo sapiens.
XX
XX WO200222660-A2.
XX
XX 21-MAR-2002.
XX
XX 10-SEP-2001; 2001MO-US026015.
XX
XX 11-SEP-2000; 2000US-00659671.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Aungl V, Zhang J, Zhao QA, Ren F;
XX Xue HJ, Yang Y, Wehman T, Dmanac RT;
XX WPI; 2002-292408/33.
XX N-PSDB; ABN32477.
XX
XX An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis.
XX
XX Example 2; SEQ ID NO 559; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat
XX stroke
XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a protein of the invention
XX
XX Sequence 665 AA;
SQ
Alignment Scores:
Pred. No.: 3,16e-287 Length: 665
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 94.22% Indels: 0
DB: 5 Gaps: 0
US-10-029-345a-108_copy_538_2532 (1-1995) x ABB97291 (1-665)
QY 1 ATGGCCATGAGATGATGGAAGTCAATGTTACTAGAGAGTGGTGCTGCTGAA 60
DB 1 MetAlHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAGTGTCTGATTAATGATAGCCGGCATTGTTGAATCAATCAATCC 120
DB 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyraenThrSer 40
QY 121 CACATTTGGAGCCATTATATCACTGCTCAAGCTTATGAAGGAGAGTTGCAACG 180
DB 41 HisIleuGluValIleAlaIleuValIleAspSerIleuValIleAspSerIleuValIleAspSer 60
QY 181 GACAAAGTGAATTAATCAAGATCAATCAAGATCAAGCAAGCAATGAAGTGAATGAT 240
DB 61 AspIleValLeuIleThrGluLeuIleGlnHisSerIleAlaIleValIleAspSer 80

QY 241 TGCAGTCAGAGGTGTAGTTTACGATCAAGCTCCCAAGATGTGCTCTCTTCA 300
DB 81 CysSerGlnIleValIleValIleAspGlnSerSerGlnAspValAlaSerLeuSer 100
QY 301 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAAAGCTTCACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyIleGluGluIleSerPheIleValIleLeu 120
QY 361 CTTCGAGGTGGTTTCTGATGTTCTGTTGTTTCTGAGCTCTGTTGAGGAAATTC 420
DB 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuGlyIleSer 140
QY 421 ACTCTGATCCCACTGATCTTCTGAGCTGCTGAGCTTCACTGTTGAGGCAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCACTGATCTTCTGAGCTGCTGAGCTTCACTGTTGAGGCAAC 540
DB 161 ArgIleLeuProAsnLeuThrLeuGlyCysGlnArgAspValIleuAsnIleGlyProThr 180
QY 541 CAGCAGAAATGGAGTGTATGTTAAATGCCAGCTATACCTGTCGAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyTyraValLeuAlaAsnAlaSerAsnThrCysProIleProAspPhe 200
QY 601 ATCCCGGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 201 IleProGlnSerIlePheLeuArgValProValAsnAspSerPheCysGlyIleLeu 220
QY 661 CCGTGTGGAACAATCAATGATGATTCATGAGAAAGCAAAAGCTTCAATGATGATG 720
DB 221 ProThrLeuAspIleSerValIleAspPheIleGluValAlaIleAsnArgIleCysVal 240
QY 721 CTAGTCACTGTTTAACTGAGATCTCCGCTCCGCAATGCTTATGCTTATGATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
QY 781 AAGAGATGAGCATGCTTCTTAAATGAGTGAAGCTTCAAGTTTGTGAAGAAAGAACTACT 840
DB 261 LeuArgMetAspPheSerLeuAspGlnAlaIleTyraPheValIleGluValAspProThr 280
QY 841 ATATCTCCAAATCTCAATTTCTGAGCAACTCTGAGTATGAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnIleLeuGlyIleLeuIleAspTyraIleValIleValIleVal 300
QY 901 CAGACTGAGCATCAGAGGCAAGCAAGCAAACTCAAGCTGCTGCACTGAGAAAGCAAT 960
DB 301 GlnThrGlyAlaSerGlyProIleSerIleValLeuIleHisIleGluIleValProAsn 320
QY 961 GAACCTGTCCCTGCTGCTCAAGAGGTGAGCAAGAAAGGAGACCGCCCTCACTCAACCC 1020
DB 321 GlnProValProAlaIleValIleSerGluGlyIleValIleValIleValIleValIleVal 340
QY 1021 TGTGCGGAGCTGCTGCTCAAGAGGAGGAGCAAGGAGGAGGAGGAGGAGGAGGAGG 1080
DB 341 CysAlaAspSerAlaThrSerGluAlaIleGlyIleGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGCGTCCCAAGCTGAGCGCTGCTGTTAGAGAGACCGCGTGTATCAGGCGCTTC 1140
DB 361 ProSerValProSerValGlnProSerLeuLeuGluIleAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGCACTGCTCCGAGAGAGGCTGAGAGAGAGCAATTAAGTCAAGCTTCTTC 1200
DB 381 SerGlyLeuHisIleSerAlaAspArgLeuGluIleAspSerAlaIleValIleValIleVal 400
QY 1201 TCTCTGATATCAATCAATTCATTAATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 401 SerLeuAspIleIleSerValIleSerTyraSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
QY 1261 TCTCATCAGAAAGATGCTTGAATATCAAACTTCACTTCACTGATGAGGAGCAAC 1320
DB 421 SerSerSerGluAspAlaLeuGluIleTyraIleValProSerThrThrLeuAspGlyThrAsn 440

QY 1321 AAGCTATGCAAGTTCCTCCCTGTTCAGGAAGTATTCGAGCAAGACTCTCCGAAACAGTCTCT 1380
DB 441 LysLeuCyGlnPheSerProValGlnLeuSerGlnGlnThrProGlnThrSerPro 460
QY 1381 GATTAAGAGGAAGCCAGCATCTCCCAAGAGCTGAGACCGCCAGGCTTCAGACAGCG 1440
DB 461 AspLysGlnGlnAlaSerIleProLysLeuGlnThrIleArgProSerPheSerGln 480
QY 1441 AGCAAGCATGATTCGCTGAGAAACAGCAAGCGAGCTGGCCGCCCAAGAGTCCCTTTTA 1500
DB 481 SerLysArgLeuHisSerValArgThrSerSerSerIleThrAlaGlnArgSerLeu 500
QY 1501 TCTCAGCTGCAATGCAAGTGGAGCGTGGAGAGCAATTCACACAGCTTCCTTTTGGC 1560
DB 501 SerProLeuHisArgSerGlnSerValGlnAspAsnIleHisThrPheLeuHis 520
QY 1561 CTTTCCACCAAGCCAGCAAGCTCAGCAAGTCTGGCTGGCTGGCTTAAAGGCTGGCAG 1620
DB 521 LeuSerThrSerGlnGlnIleuThrLysSerAlaGlyLeuGlyLeuLysGlyThrHis 540
QY 1621 TCGGATATCTTTGGCCCCCAGACCTTACCCCTTCCTGACAGCAAGCTGGATTTGGC 1680
DB 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrIlePheAla 560
QY 1681 ACAAGTCTCTCAACTTCTACTCTGCTCAGGCAATTCAGAGCAAGTCCAGTTACTCT 1740
DB 561 ThrGlnSerSerHisPheThrSerAlaSerAlaIleIleIleIleIleIleIleIleIle 580
QY 1741 GCTTACAGTGCAGCAGCTGCTCCCACTTGGCGAGACCAAGTCTATTTGTGGCAGAGCG 1800
DB 581 AlaIle 1860
QY 1801 CAGAACCCAGTACAGAGCTGCTGCGCGAGAGTGGAGTGAAGAGAGAGAGAGAGAGAG 1860
DB 601 GlnLysProSerAspArgAlaAspSerArgPheSerIleIleIleIleIleIleIleIleIle 620
QY 1861 AAGCAGTTTAAACCGAGAAGCTGCCAATATGGAATTTGGAGAGCATCATGTCAGAGAAC 1920
DB 621 LysGlnPheLysArgPheSerGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGGTACCGGAAAGAGCTGGGAGAAAGTGGGAGTCAAGTCAAGTTTCCGGGAGCATGAA 1980
DB 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerPheSerIleSerMetGln 660
QY 1981 ATCAATTAGAGTCTCC 1995
DB 661 IleIleGlnValSer 665

RESULT 11
ADA54744
ID ADA54744 standard; protein; 665 AA.
XX
AC ADA54744;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 2312.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane protein; cancer;
KW Inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PR 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
XX 24-JAN-2002; 2002US-0350435P.
XX

PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y,
XX
XX WPI: 2003-395539/38.
DR
DR N-PSDB; ADA53105.
XX
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 2312; 205bp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
XX inflammatory diseases, osteoporosis or neurological disease.
SQ Sequence 665 AA:

Alignment Scores:
Pred. No.: 1.28e-286
Score: 3399.00
Percent Similarity: 99.70%
Best Local Similarity: 99.55%
Query Match: 94.02%
DB: 6
Gaps: 0

US-10-029-345A-108_COPY_538_2532 (1-1995) x ADA54744 (1-665).

QY 1 ATGCCCATGAGATGATGGAATCTCAATTTGTATCTGAGAGTGGTGGCTCTGCTGAA 60
DB 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnLysValAlaLeuLeuGln 20
QY 61 AGTGGACGGAAGAAAGCTGCTTAATTGATACCGGCAATTTGGAAATACATACATCC 120
DB 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnThrSer 40
QY 121 CACATTTTGAAGCATTAATATCAACTGCTCCAGCTTAATGAGGAAAGGTTGCAACAG 180
DB 41 HisIleLeuGlnAlaIleAsnIleAsnLysSerLysLeuMetLysArgArgLeuGln 60
QY 181 GACAAAGTTAATTACAGAGCTCATTCAGCATTCAGCGAAACATPAGGTTGACATTGAT 240
DB 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATTTACGATCAAGCTCCCAAGAGTTGGCTCTCTCTTCA 300
DB 81 CysSerGlnLysValValIleValIleValIleValIleValIleValIleValIleVal 360
QY 301 GACTGTTTCTCAGCTGACTCTTGGGTTAACTGAGAGAGCTTCAACTGTTCACCTG 360
DB 101 AspCysPheLeuThrValIleLeuLeuLysLeuGlnLysSerPheAsnSerValHisLeu 420
QY 361 CTTGCAAGTGGGTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlyPheSerArgCysPheProGlyLeuLysGlnGlyLysSer 140
QY 421 ACTCTAGTCCCTAGCTGATTTCTCAGCTTGTCTTACTCTGTGCAACATGGGCCAACC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTCCCAATCTTTATCTTGCTGCGCAGAGAGATGCTTCAACAAGAGCTGATA 540
DB 161 LeuIleLeuProAsnLeuIleLysGlnIleArgAspValIleAsnLysGlnLeuMet 180
QY 541 CAGCAGAAATGGATGGTATGTTAAATGCGCACTATACCTGTGCAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200

QY 601 ATCCCGAGCTGATTTCTGCGGTGCTGGAATGACAGCTTTGTGAGAAATTTTG 660
 Db 201 TLeProGluSerHisPheLeuAlaGValProValAsnAspSerPheCysGluValSLeu 220
 QY 661 CCGTGGTGGCAAAATCAGTAGATTTTCATTTAGAAAAGCAAAAGCTTCATATGATGTT 720
 Db 221 ProTriPheAspIysSerValAspPheIleGluValAlaValAlaSerAsnGlyCysVal 240
 QY 721 CTATGCACTGTTTACGTGGATCTCCCGCTCCGCCACCATCGCTATCGCTCATCATCAG 780
 Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaValIleMet 260
 QY 781 AAGAGATGAGCAATGCTTATGATGAACTTACAGATTGTGAAAAGAAAAGAACTACT 840
 Db 261 LysArgMetAspPheSerLeuAspGluAlaValArgPheValIlySgluValArgProThr 280
 QY 841 ATATCTCAAACTTCATTTTCTGGGCCAACTCTGACTATGAGAAGAAATTAAGAAC 900
 Db 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIlyGluValSlySileValAsn 300
 QY 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTGGAGAAACCAAT 960
 Db 301 GlnThrGlyAlaSerGlyProIysSerIlyLeuIlyLeuLeuHisLeuGluIlyProAsn 320
 QY 961 GAACCTGCTCCTGCTGCTCAGAGGCTGAGCAGAAAAGAGAGAGCCCGCTCACTCCACC 1020
 Db 321 GluProValProAlaValSerGluGlyGlyGlnIlySerIlyThrProLeuSerProPro 340
 QY 1021 TGTGCGACTGCTGCTACTCAGAGAGCAGAGCAAAAGCCCGCTGCATCCGAGCGTG 1080
 Db 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 QY 1081 CCCAGCGTCCAGCCGCTGAGCCGCTGTTAGAGAGCAGCCCGCTGTATCAGCGCTC 1140
 Db 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 QY 1141 AGTGGCTGACCTGCTCCCGAGAGCAGGCTGGAAGACAGCAATAGCTCAAGCTTCCTC 1200
 Db 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIlySleuIlyAspSerPhe 400
 QY 1201 TCTCTGATATCAATATCATTTTATTCAGCAGCATGAGCATCTTACATGCTTC 1260
 Db 401 SerLeuAspIleIlySerValSerIlySerHisSerMetAlaAlaSerLeuHisGlyPhe 420
 QY 1261 TCTCATCAGAGATGCTTTGAAATACATACTCAAACTTTCATCTGATGGAGCAAC 1320
 Db 421 SerSerSerGluAspAlaLeuGluIlyValIlyValProSerThrThrLeuAspGlyThrAsn 440
 QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATCGAGAGAGACTCCGCAACCATGCTC 1380
 Db 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
 QY 1381 GATTAAGAGAGAGCAGCATCCCGCAAGAGCTGAGACCGCCAGGCTTCAGACAGCAG 1440
 Db 461 AspIlyGluGluAlaSerIleProIlySlyPheGlnThrAlaArgProSerAspSerGln 480
 QY 1441 AGCAAGCATTTGATTCGATCAGAACCCAGAGCATGAGCACCGCCCAAGAGTCCCTTTA 1500
 Db 481 SerIlyArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 QY 1501 TCTCCACTGATGGAAGTGGAGCGTGGAGAGCAATTAACAACAGCTTCCTTTTCGGC 1560
 Db 501 SerProLeuHisArgSerGlySerValGluAspAsnIlyHisThrSerPheLeuPheGly 520
 QY 1561 CTTTCCACAGCAGCAGCACTCACAAGAGTGTGTGGCTGTGGGCTTAAGGCTGAC 1620
 Db 521 LeuSerThrSerGlnGlnHisLeuThrIlySerAlaGlyLeuGlyLeuIlySgluValPhe 540
 QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTGATTTTGGC 1680
 Db 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerIlyPyrPheAla 560

QY 1681 ACAGAGTCTCACTTCTACTCTGCTCAGCCATCTACAGAGGAGGAGTCCAGTACTCT 1740
 Db 561 ThrGluSerSerHisPheIlySerAlaSerAlaIleIlyGlyIlySerAlaSerIlySer 580
 QY 1741 GCCTACAGCTGACAGCCAGCTGCCACTTCCGAGAGCAAAAGTCTATTTCTGTGCGACGCG 1800
 Db 581 AlaIlySerCysSerGlnLeuProThrCysGlyAspGlnValIlySerValArgArg 600
 QY 1801 CAGAACCAAGTACAGACTGACTGCGCGCGAGAGTGTGCATCAAGAGCCCTTTGA 1860
 Db 601 GlnIlyProSerAspArgAlaAspSerArgSerTrpHisGluIlyLeuSerProPheGlu 620
 QY 1861 AAGCAGTTTAAAGCAGAGTCCCAATGAAATTTGAGAGAGCATGTGACAGAAC 1920
 Db 621 LysGlnPheIlyArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
 QY 1921 AGGTCACGGAGAAAGCTGGGAGAAAGTGGCACTGACTTATGCTTTTGGGCGACATGAA 1980
 Db 641 ArgSerArgGluGluLeuGlyIlyValGlySerGlnSerSerPheSerGlySerMetGlu 660
 QY 1981 ATCATTTAGAGTCTCC 1995
 Db 661 IleIleGluValSer 665
 RESULT 12
 AAU79161
 ID AAU79161 standard; protein; 665 AA.
 AC AAU79161;
 XX 02-JUL-2002 (first entry)
 DT Human dual-specificity phosphatase-3 (DSP-16) mutant protein #1.
 DE
 XX Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; mutant; mutcin.
 XX
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 213 /note= "Wild-type Asp substituted by Ala"
 FT
 PN MO200226997-A2.
 XX 04-APR-2002.
 PD 25-SEP-2001; 2001WO-US030124.
 PF 26-SEP-2000; 2000US-0235487P.
 PR (CEPT-) CEPTYR INC.
 PA Luche RM, Wei B;
 XX MPI; 2002-315802/35.
 DR
 XX New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX
 PS Claim 46; Page; 87pp; English.
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16

CC activity, for modulation of a proliferative response in a cell, survival
 CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC but is derived from the wild-type human DSP-16 (AAU79156) protein given
 CC in figure 2 of the specification
 CC XX

Sequence 665 AA:

Alignment Scores:
 Pred. No.: 1,57e-286
 Score: 3398.00
 Percent Similarity: 99.70%
 Best Local Similarity: 99.55%
 Query Match: 94.00%
 DB: 5
 Length: 665
 Matches: 662
 Conservative: 1
 Mismatches: 2
 Indels: 0
 Gaps: 0

US-10-029-345A-108_COPY_538_2532 (1-1995) x AAU79161 (1-665)

QY 1 ATGGCCCATGAGATGATGGAATCAATTTGTTACTGAGAGGTGGCTGTGAGAA 60
 DB 1 MetAlaHisGluMetCileGlyThrGlnIleValIThrGluArgLeuValAlaLeuLeuGln 20
 QY 61 AGTGAACGGAAAAAGTCTGTAATTGATGCGCGGCATTGTGGAAATCAATACATCC 120
 DB 21 SerGlyThrGlnIuysValLeuLeuIleAspSerArgProPheValGlnIuysThrSer 40
 QY 121 CACATTTTGGAAAGCATTAATATCACTGCTCCAGCTTATGAGCGAAAGTTGCAACAG 180
 DB 41 HisIleLeuGlnIuAlaIleAsnIleAsnCysSerIysLeuMetIysArgIleuGln 60
 QY 181 GACAAAGTTTAATTACAGAGCTCATCAGCATTCAGCGAAACATAGAGTTTGAATTGAT 240
 DB 61 AspIysValLeuIleThrGlnIleuIleGlnHisSerAlaIysHisIysValAspIleAsp 80
 QY 241 TGCAGTCAGAAAGGTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
 DB 81 CysSerGlnIysValIValIValIYrAspGlnSerSerGlnAspValAlaSerIleuSer 100
 QY 301 GACTGTTTTCACCTGTAATCTTGGGTAATCGAGAAAGAGTTCACTTCACTTCACTG 360
 DB 101 AspCysPheLeuThrValIleuLeuGlnIysLeuGlnIuysSerPheAsnSerValHisIleu 120
 QY 361 CTTCAGAGTGGGTTTGTGCTGCTCTGTTGTTCCCTGCTGCTGTGAAGAAATCC 420
 DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlyIysSer 140
 QY 421 ACTCTAGTCCCTTACGATTTCTAGGCTGCTTACCTGTTGGCAACATTTGGCCCAAC 480
 DB 141 ThrLeuValProIleCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 QY 481 CSAATCTCCCAATCTTATTTGGCTGCGCAAGAGATGCTCTCAACAAAGAGCTGATA 540
 DB 161 ArgIleLeuProAsnLeuIysLeuGlnIysGlnIArgAspValIleuAsnIysGlnIuMet 180
 QY 541 CAGCGAATGGAGTTGATTATGTGTTAAATCCAGCTATACCTGTCCAAAGCTTGACTTT 600
 DB 181 GlnGlnIleuGlnIleGlyIysValIleuAsnAlaSerAsnThrCysProIysProAspPhe 200
 QY 601 ATCCCGAGTGCATTTCTGGCGTGTGCTGTGATGACAGCTTTTGTGAGAAATTTTG 660
 DB 201 IleProGlnIserHisPheIleuArgValIProValAsnAlaSerPheCysGlnIysIleu 220
 QY 661 CCGTGGTTTGAACAATAGATGATTTGATTTGAGAAAGCAAAAGCTCCCAATGAGATGTT 720
 DB 221 ProIleuAspIysSerValAspPheIleGlnIuysAlaIysAlaSerAsnIysCysVal 240

QY 721 CTATGCACTGTTTAGCTGGATCTCCGCTCCGCAACATGCTATCGCTTACATATG 780
 DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
 QY 781 AAGAGATGGAATGCTCTTTTACATGAAAGCTTACAGATTTTGTGAAGAAAAAGCACTACT 840
 DB 261 IysArgMetAspMetSerLeuAspGlnIuAlaIysArgPheValIysGlnIuysArgProThr 280
 QY 841 ATATCTCAAACTTCAATTTTCTGGCCCACTCTGTGACATATGAGAAAGAAATTAAGAC 900
 DB 281 IleSerProAsnPheAsnPheLeuGlnIleuLeuAspIysGlnIuysIleIysAsn 300
 QY 901 CAGCTGAGATGAGAGCGCCCAAGAAAGCAAACTCAAGCTGTGACCTTGAGAAAGCCAAAT 960
 DB 301 GlnThrGlyAlaSerGlyProIysSerIysLeuIysLeuHisIleuGlnIuysProAsn 320
 QY 961 GAACTGTCTGCTGTCTGAGAGGTGAGCAGAAAGAGAGAGCGCCCTCATGTCACCC 1020
 DB 321 GlnProValIProAlaValSerGlnIuysGlnIuysSerGlnIuysProIleuSerPro 340
 QY 1021 TGTGCGCACTGTGTAATCTGACAGAGCAGAGCAAAAGCCCGTCATCCGCGCAGCGTG 1080
 DB 341 CysAlaAspSerAlaThrSerGlnIuAlaIleGlnIArgProValHisIProAlaSerVal 360
 QY 1081 CCAAGCTGCGCCAGCGTGCAGCGGTGATGAGAGCAGCGCGCTGTACAGCGCTC 1140
 DB 361 ProSerValProSerValGlnProSerLeuLeuGlnIuysSerProLeuValGlnAlaIleu 380
 QY 1141 AGTGGGTGACCTGTCGCGCAGACAGAGCTGGAAGCAGACAAATTAAGCTTACGCTTCT 1200
 DB 381 SerGlyLeuHisIleuSerAlaAspArgLeuGlnIuysSerAsnIysLeuIysArgSerPhe 400
 QY 1201 TCTCTGATATCAATCAATTCATATTCAGCCAGCATGGAGAGATGCTTACATGCTTC 1260
 DB 401 SerLeuAspIleIysSerValSerIysSerIleSerIleAlaIleSerIleHisIleGlyPhe 420
 QY 1261 TCCATCTCAAGAAAGCTTTGGAATATCAACAACTTCCACTACTGTGATGGAGCAAC 1320
 DB 421 SerSerSerGlnAspAlaLeuGlnIuysIysIysProSerThrThrLeuAspGlyThrAsn 440
 QY 1321 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATCGAGAGAGACTCCGGAACACTCT 1380
 DB 441 IysIleuCysGlnPheSerProValGlnGlnIleuSerGlnIuysProGlnIuysSerPro 460
 QY 1381 GATAAGAGAAAGCAGATCCCGCAAGACTGCAACCGCGAGGCTTCAAGACGCCAG 1440
 DB 461 AspIysGlnIuAlaSerIleProIysIysLeuGlnIThrAlaArgProSerAspSerGln 480
 QY 1441 AGCAAGCATTTGCAATTCGATCAAGAACAGAGCAGAGCGCCAGAGGTCCTTTTA 1500
 DB 481 SerIysArgIleuHisSerValArgIleIserSerSerGlyThrAlaGlnIuysSerLeu 500
 QY 1501 TCTCAGTCGATGAGTGGAGAGCGTGGAGAGCAATTAACACACAGCTTCTTTTCCGC 1560
 DB 501 SerProIleuHisArgSerGlnIysSerValGlnAspAsnIuysIleThrSerPheIleuPheGly 520
 QY 1561 CTTTTCACACAGCAGAGACACTCAAGAACTGTGCTGCGCTTAAAGCTGGCAGC 1620
 DB 521 LeuSerThrSerGlnIleHisIleuThrIysSerAlaGlyLeuGlnIuysIysGlyIleHis 540
 QY 1621 TCGGATATCTTGGCCCCCGAGACTTACCTCTGACAGCAGAGCTGTATTTTGGC 1680
 DB 541 SerAspIleuAlaProGlnIleThrIleThrProSerLeuIleThrSerSerIlePheAla 560
 QY 1681 ACAAGTCTTCACTTACTTGTGCTCAGCACTTACAGGAGAGAGAGAGTCAAGTATCTT 1740
 DB 561 ThrGlnIserSerHisPheIysSerAlaSerAlaIleIysGlnIuysSerAlaSerIysSer 580
 QY 1741 GCTACAGCTGAGAGCACTGCCACTTGGGAGAGACCAAGCTTATTTGTGCGCAGCGCG 1800
 DB 581 AlaIysIserCysSerGlnIuysProIleCysGlyAspGlnIValIysSerValArgAlaGly 600
 QY 1801 CAGAAAGCAAGTGAACAGAGCTGCTCGCGGAGCTGCGATGAAGAGAGCCCTTTGAA 1860

|||||
Db 601 GlnlyseroserapargalaaapserArArSerTrphtsGluGluSerPropheGlu 620
QY 1861 AAGCAGTTTAAAGCAGAACGTCGAATGGAATTGGAGAGCATCATGTCAAGAAC 1920
Db 621 LysGlnPheIysArGArSerCySGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
QY 1921 AGGTCAAGGAGAGCGTGGGAAAGTGGGAGTCAGTCGTTTTCGGGACAGATGGAA 1980
Db 641 ArGserArGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
QY 1981 ATCATTGAGGCTCC 1995
Db 661 IleIleGluValSer 665
RESULT 13
AAU79162
ID AAU79162 standard; protein; 665 AA.
XX
AC AAU79162;
XX
DT 02-JUL-2002 (first entry)
XX
DE Human dual-specificity phosphatase-3 (DSP-16) mutant protein #2.
XX
KW Human: dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; mutant; mutain.
XX
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 244 /note="Wild-type Cys substituted by Ser"
FT
PN MO200226997-A2.
XX
PD 04-APR-2002.
XX
PF 25-SEP-2001; 2001MO-US030124.
XX
PR 26-SEP-2000; 2000US-0235487P.
XX
PA (CEPT-) CEPTYR INC.
XX
PI Luche RM, Wei B;
XX
DR WPI; 2002-315802/35.
XX
PT New DSP-16 polypeptide, useful for identifying modulators of its
PT activity, which can be used in the treatment of disorders such as
PT Duchenne muscular dystrophy, or cancer.
PS Claim 46; Page; 87pp; English.
XX
XX The present invention relates to a new polypeptide, DSP-16, having a 665
CC amino acid sequence, given in the specification, or a variant having at
CC least 50 % identical residues, which retains the ability to
CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
CC invention can be used for identifying agents which modulate DSP-16
CC activity, for modulation of a proliferative response in a cell, survival
CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne
CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
CC cell proliferation, and cell cycle abnormalities. The present amino acid
CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)

CC mutant protein #2. Note: This sequence is not shown in the specification
CC but is derived from the wild-type human DSP-16 (AAU79162) protein given
CC in figure 2 of the specification
XX
SQ Sequence 665 AA;
Alignment Scores:
Pred. No.: 2,34e-286 Length: 665
Score: 3396.00 Matches: 662
Percent Similarity: 99.70% Conservative: 1
Best Local Similarity: 99.55% Mismatches: 2
Query Match: 93.94% Indels: 0
DB: Gaps: 5 0
US-10-029-345A-108_COPY_538_2532 (1-1995) x AAU79162 (1-665)
QY 1 ATGGCCCATGAGATGATGGAACTCAAAATTTTACTGAGAGGTGTGCTGCTGGAA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
QY 61 AGTGAACGGAAAAAGTGTCTGCTAATTGATGACCGGCAATTTGCTGAATTCATACATCC 120
Db 21 SerGlyThrGluValLeuLeuLeuIleAspSerArGProPheValGluTyraenThrSer 40
QY 121 CACATTTTGAAGCCATTATATATCACTGCTCCAACTTATGAAGGAAAGTTGCAACAG 180
Db 41 HisIleLeuGlnAlaIleLeuIleLeuIleAsnGlySerLeuLeuMetLysArgLeuGlnGln 60
QY 181 GACAAAGTGTAAATTAACAAGCTCATCAGACATTCAGCGAAACATNAGTTGACATTGAT 240
Db 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGTATGATTACGATCAAGCTCCCAAGATGTGCTCTCTCTCA 300
Db 81 CysSerGlnLysValValValValTyraeGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTCACTGACTTCTGAGGTAAACTGAAGAAGCTTCAACTCTGTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGGTGGGTTTGTGAGTTCTGCTGTTGTTTCTGAGCTGTGAGGAAATTC 420
Db 121 LeuAlaGlyGlyPheAlaGluPheSerArGysPheProGlyLeuGlyLysSer 140
QY 421 ACTCTAGTCCCTACCTGCACTTCTCAGCTGCTGTTACCTGTGCAACATGGCCAAACC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTATCTTGCTGCGCAGCGAGATGTCTCAACAAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuTyraeGlyCysGlnArgAspValLeuAsnLysGluLeuMet 180
QY 541 CAGCAGAAATGGAGATTGATTATGTGTTAAATGCCAGATATACCTGTCCAAAGCTGACTTT 600
Db 181 GlnGlnAsnGlyIleGlyTyraeValLeuAsnAlaSerArThrCysProLysProAspPhe 200
QY 601 ATCCCGGAGTCTATTCTCGGTGCTGCTGCAATGAACAGCTTTGTGAGAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleLeu 220
QY 661 CCGTGTGTGACAAATCAGATGATTTCATTGAGAAAGCAAAAGCTTCAATGATGTGTT 720
Db 221 ProThrPheAspLysSerValAspPheIleGluValAlaLysAsnGlyCysVal 240
QY 721 CTATGCTACTGTTTACCTGGAGTCTCCGCTCGCCAGCATGCTATCGCTTACATCATG 780
Db 241 LeuValHisSerLeuAlaGlyLysSerArGSerAlaThrIleAlaIleAlaTyraeMet 260
QY 781 AAGAGATGACATGCTTATGATGAAGCTTCAAGATTGTGAAAGAAAAAGCACTACT 840
Db 261 LysArgMetAspMetSerLeuAspGluAlaTyraeArgPheValIleGluLysAspProThr 280
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGATGAGAAAGATTAAGAAC 900

```

Db      641 ArgSerArgGluGluGluGlyValGlySerGlnSerSerPheSerGlySerMetGln
Qy      1961 ATCTTACGCTCC 1995
Db      661 lletleclvalSer 665
RESULT 14
ID      ABR52352
ABR52352 ABR52352 standard; protein; 665 AA.
AC      ABR52352;
XX      19-JUN-2003 (first entry)
XX      19-JUN-2003 (first entry)
DE      Protein relating to the invention SEQ ID NO: 42.
KW      antiproliferative; hepatotropic; nephrotropic; antiarthritic;
KW      antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
KW      proliferative disorder; renal failure; cardiovascular disorder;
KW      immunological disorder; arthritis; psoriasis; congenital heart defect;
KW      congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
OS      Homo sapiens.
XX      WO200257460-A2.
XX      25-JUL-2002.
XX      20-DEC-2001, 2001WO-US050459.
XX      20-DEC-2000, 2000US-025668P.
XX      30-MAR-2001, 2001US-0280186P.
XX      01-MAY-2001, 2001US-0287735P.
XX      05-JUN-2001, 2001US-0295848P.
XX      25-JUN-2001, 2001US-0300465P.
XX      (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX      Jackson DG, Feder J, Nelson T, Muntier G, Ramanathan C, Lee L;
XX      Siemens N, Bol D, Schieven G, Finger J, Todderud CG, Baasolino D;
XX      Krystek S, Mcatee P, Suchard S, Banas D;
XX      WPI; 2002-599721/64.
XX      N-PSDB; ACC60521.
XX      Novel polynucleotides encoding human phosphatase polypeptides useful in
XX      the prevention or treatment of e.g. proliferative and cardiovascular
XX      disorders.
XX      Claim 5, Fig 5; 801pp; English.
CC      The invention relates to a novel isolated nucleic acid comprising a
CC      polynucleotide having a nucleotide sequence selected from 40
CC      polynucleotides fully defined in the specification. The polynucleotide of
CC      the invention has antiproliferative, hepatotropic, nephrotropic,
CC      antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC      polynucleotide may have a use in gene therapy. A polynucleotide or
CC      polypeptide of the invention is useful for preventing, treating or
CC      also useful for a medical condition, e.g. a proliferative disorder or
CC      disorders including arthritis and psoriasis, renal failure, immunological
CC      cancer. A method of the invention is useful for diagnosing a pathological
CC      condition or susceptibility to a condition in a subject. The present
CC      sequence is used in the exemplification of the invention
XX      Sequence 665 AA;
XX
Alignment Scores:
Pred. No.: 1.16e-285 Length: 665
Score: 3388.00 Matches: 661
Percent Similarity: 99.55% Conservative: 1

```

Query Match: 93.72% Indels: 0
DB: 5 Gaps: 0
US-10-029-345a-108_copy_538_2532 (1-1995) x ABR52352 (1-665)

```
QY 1 ATGGCCATGATGATTTGGAACCTCAAAATGTTACTAGAGAGTTGGTGGCTGTCTGGAA 60
DB 1 MetAlaHisGluMetCileGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
QY 61 AGTGAACCGAAAAAGTGTCTTAATTGATAGCCGCGCATTTGGAAATCAATACATCC 120
DB 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyraAnthSer 40
QY 121 CAAATTTTGAACCATTAATATCACTGCTCAAGTTTGAAGCAAGATTGGCAACG 180
DB 41 HisIleLeuGlnIleAlaIleAsnIleAsnCySerIleuMetIleAspArgLeuGln 60
QY 181 GACAAAGTGAATTAATACAGAGCTCATCCAGCATTCAGCAAAACATAAGTTGACATTGAT 240
DB 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATGTTACGATCAAAAGCTCCCAAGATGTGCTCTCTTCA 300
DB 81 CySerGlnLysValValValTyraAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTCACTGATCTTCTGGGTTAACTGGAAAGAGCTTCACTGTTCACCTG 360
DB 101 AspCyAspHeuThrValIleuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGTGGGTTTGGTGGATTCTCTCGTTGTTTCCCTGGGCTGTGTGAAGAAATCC 420
DB 121 LeuAlaGlyIleGlyIleAsnGluPheSerArgCyPheProGlyLeuCyGlnGlyLysSer 140
QY 421 ACTTGAATCCCTACCTGACATTTCTCAGCCTTGTCTTACCTTGTGCCAACATTTGGCCAAAC 480
DB 141 ThrLeuValProThrCyHisSerGlnProCySerLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTTATCTTGGTGGCCAGCGAGATGCTCTCAACAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuTyLeuGlyCyGlnArgAspValLeuAsnLysGluLeuMet 180
QY 541 CACGAGAAATGGATGGTATGTTAAATGCGCAGTATACCTGTCCAAAGCCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyTyraIleAsnAlaSerAsnThrCySerProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGGTGGCTGTGAATGACAGCTTTGTGAAAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCyGlnLysIleLeu 220
QY 661 CCGTGGTTGACAAATAGTATGATTTCTTGAAGAAAGCAAAAGCTTCATGATGTT 720
DB 221 ProThrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCyVal 240
QY 721 CTAGTGACATGTTAGTGGATCTCCGCTCCGCCACATCGCTATCGCTCATCATG 780
DB 241 LeuValHisCySerLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260
QY 781 AAGAGATGACATGTTCTTTAGTGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaTyraArgPheValLysGlnLysValAspProThr 280
QY 841 ATATCTCCAAATCTTCAATTTTGGGCGCACTCTGACATATGAAAGAAAGATTAAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAlaTyrlGlnLysLysIleLysAsn 300
QY 901 CAGACTGAGCATCAGGGCGCAAAAGCAAACTCAAGTGTGCACTCGGAGGAGCCAAAT 960
DB 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuProLeuGlnLysProAsn 320
QY 961 GAACTGTCTCTGTCTCAGAGAGGTGACAGAAAAAGCAGAGCGCCCTCACTCCACC 1020
DB 321 GluProValProAlaValSerGlnGlyGlyGlnLysSerGlnThrProLeuSerProPro 340
```

```
QY 1021 TGTGCCACTCTGCTCACTCAGAGGAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
DB 341 CyAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGCGTCCCGCAGCGTGCAGCCGCTGTTTGAAGACAGCCCGCTGTACAGCGCTC 1140
DB 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGCACTGTCCCGCAGACAGCGTGGAAAGACAAATTAAGCTCAAGCTTCTTC 1200
DB 381 SerGlyLeuHisIleuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 400
QY 1201 TCTCTGATATCAAAATCAGTTTATATATGACCCAGATGGCAGCATCTTACATGCTTC 1260
DB 401 SerLeuAspLysSerValSerTyraSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
QY 1261 TCTCATCAGAAAGATCTTTGGAATCTCAAAACCTTCCATCTGTGATGGAGCCAAC 1320
DB 421 SerSerSerGlnAspAlaLeuGlnTyrlTyraProSerThrThrLeuAspGlyThrAsn 440
QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAGAAATATGCGAGACAGCTCCGAAACCATGCTT 1380
DB 441 LysLeuCyGlnPheSerProValGlnGlnLysSerGlnGlnThrProGlnThrSerPro 460
QY 1381 GATTAAGAGGAGGAGCCAGCATCCCGCAAGAGCTGCAGACCGCCAGGCTTCAGACAGCAG 1440
DB 461 AspLysGlnGlnLysSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
QY 1441 AGCAAGCATTTGATTCGGTCAGAAACAGCAGCAGTGGCAGCCGCGCAGAGTCCCTTTTA 1500
DB 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
QY 1501 TCTCCACTGCATGGAAGTGGAGCGTGGAGAGACAAATTAACACACGACTTCTTTTGGCC 1560
DB 501 SerProLeuHisArgSerGlySerValGlnAspAsnTyrlHisPheSerPheLeuPheGly 520
QY 1561 CTTTCCAGCAGCAGAGACCTCAGAAAGTCTGCGGCTGGGCTTAAAGGCTGGCAC 1620
DB 521 LeuSerThrSerGlnIleHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyThrHis 540
QY 1621 TCGGATATCTTGGCCCCCGCAGACCTTACCCCTTCCCTGACACAGCAGCTGGTATTTGGCC 1680
DB 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrlPheAla 560
QY 1681 ACAGAGTCTCACACTTTACTCTGCTCAACCATTAACGAGGAGCAGTGCAGTTACTCT 1740
DB 561 ThrGlnSerSerHisPheTyrlSerAlaSerAlaIleTyrlGlyLysSerAlaSerTyrlSer 580
QY 1741 GCTTACAGCTGCAGCGAGCGCCCACTTGCAGAGCAAGTCAATCTGTGGCGCAGGCGG 1800
DB 581 AlaTyrlSerCySerGlnLeuProThrCyGlyAspGlnValTyrlSerValArgArgArg 600
QY 1801 CAGAAACCAAGTGCAGAGCTGACTGCGCGCGGAGCTGGCATGAAAGAGACCCCTTTGAA 1860
DB 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrlHisGlnGlnSerProPheGln 620
QY 1861 AAGCAGTTTAAACGCAAGAGCTGCGCAATGGAATTTGAGAGAGCATCATGTAGAGAAC 1920
DB 621 LysGlnPheLysArgArgSerCyGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGGTCAAGGGAAGAGCTGGGGGAAAGTGGGAGTCACTTACAGCTTTTCCGGGACAGTGA 1980
DB 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATTTAGAGTCTCC 1995
DB 661 IleIleGlnValSer 665
```

RESULT 15
AAB20325 ID AAB20325 standard; protein; 666 AA.
XX AAB20325;

XX	29-MAY-2001	(first entry)	
DT			
XX			
DE	Human protein phosphatase and kinase protein-4.		
XX			
KM	Protein phosphatase and kinase protein, PPKR-4; human;		
KM	gastrointestinal disorder; immune system disorder; neurological disorder		
KW	cell proliferative disorder; cancer; diagnosis; therapy.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Modified-site	12	
FT		/note= "O-phosphorylated"	
FT	Modified-site	21	
FT		/note= "O-phosphorylated"	
FT	Modified-site	23	
FT		/note= "O-phosphorylated"	
FT	Modified-site	38	
FT		/note= "N-glycosylated"	
FT	Modified-site	38	
FT		/note= "N-glycosylated"	
FT	Modified-site	49	
FT		/note= "N-glycosylated"	
FT	Modified-site	72	
FT		/note= "O-phosphorylated"	
FT	Modified-site	82	
FT		/note= "O-phosphorylated"	
FT	Region	85..298	
FT		/note= "VH1-type dual specificity phosphatase signature"	
FT	Modified-site	91	
FT		/note= "O-phosphorylated"	
FT	Modified-site	190	
FT		/note= "N-glycosylated"	
FT	Modified-site	212	
FT		/note= "N-glycosylated"	
FT	Modified-site	214	
FT		/note= "O-phosphorylated"	
FT	Active-site	220..280	
FT		/note= "tyrosine specific protein phosphatase"	
FT	Region	237..278	
FT		/note= "Y phosphatase signature"	
FT	Modified-site	266	
FT		/note= "O-phosphorylated"	
FT	Modified-site	280	
FT		/note= "O-phosphorylated"	
FT	Modified-site	300	
FT		/note= "N-glycosylated"	
FT	Modified-site	369	
FT		/note= "O-phosphorylated"	
FT	Modified-site	393	
FT		/note= "O-phosphorylated"	
FT	Modified-site	421	
FT		/note= "O-phosphorylated"	
FT	Modified-site	422	
FT		/note= "O-phosphorylated"	
FT	Modified-site	434	
FT		/note= "O-phosphorylated"	
FT	Modified-site	439	
FT		/note= "O-phosphorylated"	
FT	Modified-site	468	
FT		/note= "O-phosphorylated"	
FT	Modified-site	471	
FT		/note= "O-phosphorylated"	
FT	Modified-site	479	
FT		/note= "O-phosphorylated"	
FT	Modified-site	528	
FT		/note= "O-phosphorylated"	
FT	Modified-site	590	
FT		/note= "O-phosphorylated"	
FT	Modified-site	597	
FT		/note= "O-phosphorylated"	
FT	Modified-site	605	

FT	/note= "O-phosphorylated"
FT	Modified-site 610
FT	/note= "O-phosphorylated"
FT	Modified-site 613
FT	/note= "O-phosphorylated"
FT	Modified-site 618
FT	/note= "O-phosphorylated"
FT	Modified-site 628
FT	/note= "O-phosphorylated"
FT	Modified-site 641
FT	/note= "N-glycosylated"
FT	Modified-site 643
FT	/note= "O-phosphorylated"
PX	WO200120004-A2.
PN	
PP	22-MAR-2001.
PR	14-SEP-2000; 2000WO-US025515.
PR	15-SEP-1999; 99US-015414IP.
XX	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Yue H, Tang YT, Bandman O, Hillman JL, Baughn NR, Azimzai Y, Lu DM;
DR	WPJ; 2001-244811/25.
XX	N-PSDB; AAF30479.
PT	Novel human protein phosphatase and kinase proteins for diagnosis, treatment and prevention of gastrointestinal, immune system, neurological and cell proliferative disorders.
PS	Claim 1; Page 87-88; 103pp; English.
CC	The present sequence is that of novel human protein phosphatase and kinase protein PPHK-4, as predicted from Incyte Clone ID No. 1234795CB1 (see AAF30479). Tissues that express PPHK-4 (as a fraction of total tissues expressing PPHK-4) include gastrointestinal (0.385), cardiovascular (0.154), nervous (0.154) and reproductive (0.154). Diseases or conditions associated with tissues expressing PPHK-4 (as a fraction of total tissues expressing PPHK-4) include cancer (0.692), inflammation or trauma (0.308) and cell proliferation (0.231). The encoded protein shows homology to mouse neuronal tyrosine threonine phosphatase 1. The invention provides human PPHK-1 co-11 polypeptides (see AAB20322-32) and polynucleotides (see AAF30476-86). It also provides expression vectors, host cells, antibodies, agonists and antagonists, as well as methods for diagnosing, treating or preventing disorders associated with expression of PPHK, including gastrointestinal disorders, immune system disorders, neurological disorders and cell proliferative disorders, including cancer
SQ	Sequence 666 AA:
Alignment Scores:	
Pred. No.:	6,4e-285
Score:	3379.50
Percent Similarity:	99.40%
Best Local Similarity:	99.25%
Query Match:	93.49%
DB:	Gaps: 4
US-10-029-345A-108_COPY_538_2532 (1-1995) x AAB20325 (1-666)	
OY	1 ATGGCCCATGATATTGTGAACCAATTGTTACTGAAGGTTGGCTTGTCTGGAA 60
Db	1 MetAlahsgluMetIleGlyThrGlnIleValThrGluArgLeuValAlaIleuGlu 20
OY	61 AGTGGAACGAAAAAGTCGTCTTAATGTGAAGCGCGGCCCTTGTGGATCAATATCACC 120
Db	21 SerGlyThnGluIysValIleuIleuIleAspSerArgProPheValGluIysAsnThrSer 40

OY	12	CACATTTTGGAAAGCCATTAATATCAACCTGCTCAAGCTTTATGAAAGCAAGTTGGACAAG	180
Db	41	HisIleuLeuGluAlaIleasnIleasnucyserylshleuMetLysArgLeuGlnGln	60
OY	181	GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGGAAACATMAAGTTGACATTGAT	240
Db	61	AerLysValLeuIleLeuThrGluLeuIleGlnHisSerAlaLysHisLysValAspLLeuArg	80
OY	241	TGCAGTCAAAAGGTGTGATTGTTACGATCAAAAGCTCCAAAGATGTTGGCTTCTCTTCA	300
Db	81	CysSerGlnLysValValValLysArgGlnSerSerGlnAspValAlaSerLysSerSer	100
OY	301	GACGTGTTTTCACtGTACTTCTGGGTAAACtGGAGAAAGAGTTCAACTGTGTTCACTG	360
Db	101	AspCysPheLeuThrValLeuLeuGlnLysLysLeuGlnLysSerPheAsnSerValHisLeu	120
OY	361	CTTCAGAGTGGGGTGTGCTAGTTCCTCTGTTGTTTCCCTGGAGCTCTGTGAAGAAATCC	420
Db	121	LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlySerSer	140
OY	421	ACTGTAGTCCCTACCTGCATTTTCTCAGCCTTGCTTACTCTGTTGCCAATTTGGGCAAC	480
Db	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
OY	481	CGAATTTCTCCCAACTTTATCTTGGTGGCCGACGAGATGTCTTAACAAGAAGCTGATA	540
Db	161	ArgIleLeuProAsnLeuTyrLeuGlnCysGlnAspValLeuAsnLysGluLeuMet	180
OY	541	CAGCAGATATGGGATTTGTTATGTGTTAATGGCAGCTTATACCTGTGCCAAAGCCTCACTT	600
Db	181	GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe	200
OY	601	ATCCCGAGTCTCATTTCTCTGCGTGGCTGTGAATGACAGCTTTTGTGAGAAAATTTTG	660
Db	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu	220
OY	661	CCGTGCGTTGGACAATTCAGTAGATTTCATTGAGAAAGCAAAAGCTCCAAATGATGTGT	720
Db	221	ProThrLeuAspLysSerValAspPheIleGluValAlaLysAlaSerAsnGlyCysVal	240
OY	721	CTAGTGCACGTGTTAGTGGGATCTCCGCGCTCCGCGACATGCATGCGCTCACTCATG	780
Db	241	LeuValHisCysLeuAlaGlyLysSerArgSerAlaThrIleAlaIleAlaTyrIleMet	260
OY	781	AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAGACCTACT	840
Db	261	LysArgMetAspMetSerLeuAspLysAlaTyrArgPheValLysGluLysArgProThr	280
OY	841	ATATTCCTCAAACTTAATTTTCTGGGCCCACTCTCGACATATGAGAAAGATTAAAGAC	900
Db	281	IleSerProAsnPheAsnPheLeuGlnLysLeuAspTyrGlnLysLysIleLysAsn	300
OY	901	CAGACTGGAGATCAAGGGCCAAAGAGCAAACTCAGAGCTGCGACCTGGAGAGGCAAT	960
Db	301	GlnThrGlyAlaSerGlyProLysSerLysLeuLysLysLeuHisLysGlnLysProAsn	320
OY	961	GAACCTGTCCCTGCTGCTCAGAGGGGTGAGACAGAAAGCGACGCCCTCAGTCAACC	1020
Db	321	GluProValProAlaValSerGluGlyGlyGlnLysSerGlnThrProLeuSerProPro	340
OY	1021	TGTGCGCACTGTCTACTCAGAGGACGACGAGCAAAAGCCCTGTCATCCGCGACGCTG	1080
Db	341	CysAlaAspSerAlaThrSerGlnAlaIleGlyGlnArgProValHisProAlaSerVal	360
OY	1081	CCCAAGGTGCCAGGCGTGAGCCGCTGCTTAAAGGACAGCCGCGCTGTACAGCGGCTC	1140
Db	361	ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu	380
OY	1141	AGTGGGCTGCACCTTCCGACAGACAGGCTGGAAAGACAGCAATTAAGCTCAAGGCTCTTC	1200
Db	381	SerGlyLeuHisIleLeuSerHisAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe	400
OY	1201	TCTCTGATATCAAAATCAAGTTTCATATTTACGCCAGCATGGCAGATCCTTACATGGCTTC	1260

[illegible]

Search completed: June 21, 2004, 12:34:14
Job time : 154.589 secs

the page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:26 ; Search time 15.7715 Seconds
(without alignments)
11051.549 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 1627
Sequence: 1 atgagccatgagatgattg9.....agaagattaagaccagacac 906

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=x1p
-Q=/cgn21/USPTO.spool_p/US10029345/runat_21062004_122816_4205/app_query.fasca_1.3278
-DB=pir_78 -QMT=fastcan -SUFFIX=xpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMF=prco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10029345_QCGN_1_1_78@runat_21062004_122816_4205 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	450	27.7	619	2	T15969
2	425	26.1	367	1	S24411
3	423	26.0	394	2	A56115
4	418	25.7	314	1	A57126
5	415	25.5	367	1	S29090
6	415	25.5	367	2	S52265
7	409	25.1	314	2	B57126
8	377.5	23.2	393	2	A56947
9	351	21.6	384	1	T13890
10	325	20.0	303	2	T16405
11	289	17.8	365	2	T32494
12	223.5	13.7	330	2	T39698
13	222.5	13.7	186	2	T16056
14	212.5	13.1	226	2	T21380

15	208	12.8	272	2	T18915	hypothetical prote
16	207.5	12.8	220	2	JC7885	low-molecular-mass
17	205	12.6	223	2	T49365	protein tyrosine p
18	202	12.4	354	1	S31304	protein-tyrosine-p
19	201.5	12.4	142	2	T03074	dual specificity p
20	195.5	12.0	185	1	A47196	dual specificity p
21	195	12.0	283	2	G84458	probable protein p
22	193	11.9	205	2	T49364	protein tyrosine p
23	185	11.4	807	1	S44538	probable protein-t
24	183.5	11.3	276	2	T48806	protein-tyrosine-p
25	181.5	11.2	278	2	T39517	dual-specificity M
26	179.5	11.0	866	2	F88481	protein C16a3.1 (1
27	173.5	10.7	204	2	T17802	hypothetical prote
28	171	10.5	489	1	S58725	dual specificity p
29	162	10.0	771	2	T47665	phosphatase-like p
30	160.5	9.9	580	2	T18439	hypothetical prote
31	160.5	9.9	600	2	T18446	hypothetical prote
32	157	9.6	209	1	S48459	probable dual spec
33	152	9.3	292	2	S41012	hypothetical prote
34	151.5	9.3	597	1	S43743	probable dual spec
35	150.5	9.3	169	2	T30684	probable dual spec
36	145.5	8.9	171	1	T36845	dual specificity p
37	145.5	8.9	171	2	T28522	probable dual spec
38	144.5	8.9	171	2	B72161	dual specificity p
39	144.5	8.9	171	1	Q0VZH1	dual specificity p
40	138.5	8.5	171	1	A42514	hypothetical prote
41	134	8.2	272	2	T21489	hypothetical prote
42	134	8.2	272	2	T19418	hypothetical prote
43	127.5	7.8	171	1	B47452	dual specificity p
44	126	7.7	1132	2	T49403	related to protein
45	121	7.4	928	2	S50578	hypothetical prote

ALIGNMENTS

RESULT 1

T15969
hypothetical protein F08B1.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T15969
R/Chisoe, S.
submitted to the EMBL Data Library, July 1995

A/Description: The sequence of C. elegans cosmid F08B1.
A/Reference number: Z18439

A/Accession: T15969
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
A/Residues: 1-619 <CH1>

A/Cross-references: EMBL:U23178; NID:G726421; PID:G726422; PIDN:AC46719.1; CESP:F08B1.1
A/Experimental source: strain Bristol N2.

C/Genetics:
A/Gene: CESP:F08B1.1
A/Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatch:	Indels:	Gaps:
		5.25e-32	450.00	90	34	48	6
			69.66%				
			50.56%				
			27.66%				
DB:			2			4	

US-10-029-345A-108_COPY_538_1443 (1-906) x T15969 (1-619)

QY	367	GGTGGGTTGCTAGTTCCTGCTGTTTCCTGCGCCCTGTGAAGAAATCC-----	420
DB	99	GLYGLYPheIsgInPheIsgInGlnTyPProGInIleucYsgIleuSerGlnGlyMet	118
QY	421	ACTCTAGTCCCTACCCGACATTTCTGACGCTTGTCTTA-----CCTGTGCCAACATTGGG	474
DB	119	ThrArgIleuProIleuSerIleuSerGlnProCyIleuSerGlnProThrIleuAsp--Gly	137


```

Db      312 Val 312

RESULT 6
552265 dual specificity phosphatase (EC 3.1.3.-) 1 - rat
N/Alternate names: protein-tyrosine-phosphatase Cti00; protein-tyrosine-phosphatase, not
C/Species: Rattus norvegicus (Norway rat)
C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999
C/Accession: S52265
R/Muda, M.; Schlegel, W.; Arkinsteil, S.
submitted to the EMBL Data Library, January 1995
A/Description: Pathways regulating Cti00 gene expression in pituitary cells.
A/Reference number: S52265
A/Accession: S52265
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-367 <MUD>
A/Cross-references: EMBL:X84004; NID:9642264; PIDN:CAA58828.1; PID:9642265
C/Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
C/Keyword: phosphoprotein; phosphoric monoester hydrolase
F/181-312/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
F/258/Active site: Cys (phosphocysteine intermediate) #status predicted
F/264/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:
Pred. No.: 6,7e-29 Length: 367
Score: 415.00 Matches: 104
Percent Similarity: 49.67% Conservative: 48
Best Local Similarity: 33.99% Mismatches: 126
Query Match: 25.51% Indels: 28
DB: Gaps: 6

US-10-029-345a-108_COPY_538_1443 (1-906) x S52265 (1-367)
QY 43 TTGTGGCTGCTGCGAAAGTGAACGAAAGATGCTGCTAATTGATCGCGCATTT 102
Db 13 LeuArgAlaLeuLeuArgGluArgAlaAlaGlnCysLeuLeuLeuAspCysArgSerPhe 32
QY 103 GTGAATACAAATCATCCACATTTTGAAGCCATTAAATCAATGCTCCAAAGCTTANG 162
Db 33 PheAlaPheAsnAlaGlnHisIleValGlnSerValAsnValArgPheSerThrIleVal 52
QY 163 AACCGAAGCTTGCA-----CAGACAAAGTGTAAATTAACAGCTCATC 207
Db 53 ArgArgArgAlaIleArgGlyAlaMetGlyLeuGlnHisIleValProAsnThrGlnLeu-- 71
QY 208 CAGCATTCAGCAACATACATGATTCAGATTCAGATTCAGCAAGAGTTGATTCAGAT 267
Db 72 -----ArgGlyArgLeuLeuAlaGlyAlaIleValHisAlaValValLeuLeuAsp 87
QY 268 CAAGACTCCCAAGATGTTGCTCTCTCTTCACAGCTGTTTTCATGCTGATTCGGGT 327
Db 88 GluArgSerAlaAlaLeuAspGlyAlaIleArgAspGlyThrLeuAlaLeuAlaIleVal 107
QY 328 AAATCG-----GAGAAAGACTTCACTGTTGCTGCTGCTGAGGTGGCTGAG 381
Db 108 AlaLeuCysArgGlnAlaIleArgSerThrGlnValPhePheLeuGlnGlyIleVal 127
QY 382 TTCTCTGTTGTTCCCTGCTGCTGCTGGAAGAAA----- 417
Db 128 PheSerAlaSerCysProGlnLeuCysSerIleGlnSerThrProMetGlyLeuSerLeu 147
QY 418 -----TCCACTGTAAGTCCCTACCTGATTCCTCAGCTTCCTTACCTGTCGAACATT 471
Db 148 ProLeuSerThrSerValProAspSerAlaGlnSerGlyCysSerSerCysSerThrPro 167
QY 472 -----GGGCGCAACCGCAATCTTCCCAATTTTATCTTCTGCTGCCAGCA 516
Db 168 LeuTyrAspGlnGlyIleProValGlnIleLeuSerPheLeuTyrLeuGlySerAlaTyr 187
QY 517 GATGTCCTCAACAGAGCTGATGATGACAGAAATGGATGGATGATGATGATGATGATGATG 576
Db 517 GATGTCCTCAACAGAGCTGATGATGACAGAAATGGATGGATGATGATGATGATGATGATG

```

```

Db      188 HisAlaSerArgIleAspMetLeuAspAlaLeuGlyIleThrAlaLeuIleAsnValSer 207
QY 577 TATACCTGTCCAAAGCTTATATCCCGAGTCTCATTTCTCGCTGCGCTGTAAT 636
Db 208 AlaAsnCysProAsn--HisPheGlnGlyHisIleTyrGlnTyrSerIleProValGlu 226
QY 637 GACAGCTTTGTGAGAAATTTTGGCGGTGGTGCAGCAATCAGATGATTCATGAGAAA 696
Db 227 AspAsnHisIleValAlaAspIleSerSerThrPheAsnGlnAlaIleAspPheIleAspSer 246
QY 697 GCAAAAGCCCTCAAGATGATGTTCTTATGTCAGCTGTTTACCTGGAGTCCCGCTCGCC 756
Db 247 IleValAspAlaGlyArgValPheValHisCysGlnAlaGlyIleSerArgSerAla 266
QY 757 ACCATGCGTATCCCTTACATCATGAAAGAGATGACATGCTTTATGTAAGCTTACAGA 816
Db 267 ThrIleCysLeuAlaTyrIleMetArgThrAsnArgValIleLeuAspGlnAlaPheGlu 286
QY 817 TTGTGAAGAAAGAAAGACTTATATCTGCAAACTTCAATTTTCTGGGCGCAACTCTG 876
Db 287 PheValIleGlnArgArgSerIleIleSerProAsnPheSerPheMetGlyIleLeu 306
QY 877 GACTATGAGAAAGATTT 894
Db 307 GlnPheGlnSerGlnVal 312

RESULT 7
B57126 dual specificity phosphatase (EC 3.1.3.-) 2 - mouse
N/Alternate names: mitogen-induced nuclear protein-tyrosine-phosphatase PAC-1
C/Species: Mus musculus (house mouse)
C/Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 24-Apr-1998
C/Accession: B57126
R/Rohan, P.J.; Davis, P.; Moskaluk, C.A.; Kearns, M.; Krutzsch, H.; Siebenlist, U.; Kelly,
Science 259, 1763-1766, 1993
A/Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
A/Reference number: A57126; MUID:93206122; PMID:7681221
A/Accession: B57126
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-314 <ROH>
A/Cross-references: GB:111330
C/Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
C/Keyword: nucleus; phosphoprotein; phosphoric monoester hydrolase
F/180-311/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
F/257/Active site: Cys (phosphocysteine intermediate) #status predicted
F/263/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:
Pred. No.: 2,27e-28 Length: 314
Score: 409.00 Matches: 107
Percent Similarity: 51.18% Conservative: 45
Best Local Similarity: 36.03% Mismatches: 113
Query Match: 25.14% Indels: 32
DB: Gaps: 7

US-10-029-345a-108_COPY_538_1443 (1-906) x B57126 (1-314)
QY 70 GAAAAGTCTGCTGTAATGATGACCGCATTTGTGAATACAAATACATCCACATTTTG 129
Db 25 GluArgThrLeuLeuLeuAspCysArgProPheLeuAlaPheCysArgSerHisValArg 44
QY 130 GAAGCATTAATATCACTGCTCCAAAGTTTATGACGGAAG----- 171
Db 45 AlaIleArgProValProThrAsnAlaLeuLeuArgArgAlaProGlyThrProAla 64
QY 172 -----TTGCAACAGCAAGATGTTAATTAACAGAGCTCATCCAGATTCA 216
Db 65 AlaAlaLeuAlaCysLeuLeuProAspArgAlaLeuArgAlaIleArgLeuGlyArgGly 84
QY 217 GCGAAACATTAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 276
Db 85 Leu-----AlaArgAlaValValLeuLeuAspGlnSerSer 95

```

Query: 277 CAGATGTTGCCTCTCTCTCTTCAGACTGTTTTCTACGTACTTCTGGATAAATTGAG 336
Db 96 AlAservAlThrGluLeuProProaspBglyProAlHisIleuLeuLeuAlaIalaGln 115
QY 337 AAGAAGCTTC-----AACTGTTCAACTGCTTGTCAGAGTGGATTTCTGAGTTTC 384
Db 116 HisGluMetArgIylgylProthrTrnValCysPheLeuArgIyglYlPheIysSerPhe 135
QY 385 TCTCGTTGTTTCCCTGGCTCTGT--GAAGAAAAATCCACTGTAAGCTCT-- 432
Db 136 GlnThrTyCysProaspLeuCysserGlnAlaProAlaGlnAlaIalaLeuProAlaIyl 155
QY 433 -----ACCTGATTTCTCAGCCTTGCTTACTCTGTTGCCAACATT--GGGCCAACCCGA 483
Db 156 AlaGluAsnSerAsnSerAspProArgvalProIleTyraBpgInglyIylProValGlu 175
QY 484 ATTTTCCCCAATCTTTATCTTGCTGCTGCCAGAGATGCTCCACAAGAGAGCTGATACAG 543
Db 176 IleLeuProTyrlzLeuTyrlzLeuGlyserCysasnHisSerSerAspLeugInglyLeuGln 195
QY 544 CAGATGGAATGGTTATGTGTAAATGCCAGCATACCTGTCACAAAGCCTGACTTATC 603
Db 196 AlaCysglyIleThrAlaValaIleuAsnValSerAlaSerCysProasn--HisPheGlu 214
QY 604 CCCGAGCTCATTTCCGCGCTGCTGCTGATGACAGACTTTGTGAGAAAATTTTGGCCG 663
Db 215 GlyLeuPheHisTyrlzYlsSerIleProValIGluAspAsngImetValdIuleSerAla 234
QY 664 TGTTTGGACAAATCACTAGATTTCATTGAGAAAGCAAAGCCCTCAATGATGTGTTCTA 723
Db 235 TrpPheGlnGluAlaIleSerPheIleAspSerValIysAsnSerGlyIylArgValIeu 254
QY 724 GTGCACTGTTTAGCTGGAGATCTCCGCTCCGCCACCATGCTGATGCCCTACATCATGAG 783
Db 255 ValHisCysGlnAlaGlyIleSerarSeralatrrIleCysLeuAlaTyIleuIleGln 274
QY 784 AGGATGACATGCTCTTATGATGACATGACATTTGAGAAAGAAAAAGACTACTATTA 843
Db 275 SerHisAtgValaArgLeuAspGluAlaIleAspPheValIysGlnIatgArGlyValIle 894
QY 844 TCTCCAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATT 894
Db 295 SerProaspPheSerPheMetGlyGlnIleuLeuGlnIleuGlnIleuGlnIleuGlnIle 311

RESULT 8

A: dual specificity phosphatase (EC 3.1.3.-) HVH2 - rat
N: Alternate names: mitogen-activated protein kinase phosphatase 2
C: Species: Rattus norvegicus (Norway rat)
C: Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-1999
C: Accession: A56947
R: Mistr-Press, A.; Rim, C.S.; Yao, H.; Robertson, M.S.; Stork, P.J.S.
U: J. Biol. Chem. 270, 14587-14596, 1995
A: Title: A novel mitogen-activated protein kinase phosphatase. Structure, expression, at
A: Reference number: A56947; PMID:95301550; PMID:7782322
A: Accession: A56947
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-393 <MIS->
C: Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity
C: Keywords: phosphoprotein; phosphoric monooester hydrolase
F: 202-333/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>
F: 279/Active site: Cys [phosphocysteine intermediate] #status predicted
F: 285/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:
1.	49e-25	393	377.50	95	54	124
Best Local Similarity:	48.69%					
Query Match:	31.05%					
B:	23.20%					

[illegible]

RESULT 9
138890
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 5 - human
N:Alternate names: dual specificity phosphatase B23; dual-specificity phosphatase hvh-3;
C:Species: Homo sapiens (man)
C>Date: 16-Feb-1996 #sequence_revision 01-Mar-1996 #next_change 11-Jun-1999

A:Accession: J33698
A:Status: preliminary, translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-330 <MOO>
A:Cross-references: EMBL:AL109652; PIDB:CAB51765.1; GSPDB:GN00067

A:Experimental source: strain 972h-; cosmid c17a3
C:Gene: p1041
A:Map position: 2
Alignment Scores:
Pred. No.: 8 01e-12 Length: 330
Score: 223.50 Matches: 53
Percent Similarity: 52.59% Conservative: 18
Best Local Similarity: 39.26% Mismatches: 63
Query Match: 13.74% Indels: 1
DB: 2 Gaps: 1
US-10-029-345A-108_COPY_538_1443 (1-906) x T39698 (1-330)
QY 478 ACCGGAATTCCTCCCAATCTTATCTTGCTGCCAGGAGATGCTCAACAGAGCTG 537
Db 48 SerGluIleSerIysAenLeuTyrIleSerSerTrpYthrAlaSerGluLeuValSer 67
QY 538 ATACAGCAGAAATGGGATGGTATGTATGTTAAATGCCAGCTATACCTGTCCAAAGCTGAC 597
Db 68 ThrSerAspLysGlyIleAspTyrThrLeuSerAlaMetSerIleAsnProAsnLeuSer 87
QY 598 TTATCCCGGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAGAAATT 657
Db 88 ---ValProGluGlnGlnHisLeuTrpLeuGlnIleGluAspSerSerSerGlnAsnIle 106
QY 658 TTGCGGTGGTGGACAAATGAGATGATTCATTGAGAAAGCAAAAGCTTCAGATGT 717
Db 107 LeuGlnTyrRheGluLysSerAsnLysPheIleAlaPheAlaLeuSerLysAsnAlaLys 126
QY 718 GTTCTAGTCACTGTTTAACTGGGATCTCCGCTCCGCCCAATCGCTATCGCTACATC 777
Db 127 ValLeuValHisCysPheAlaGlyIleSerArgSerValThrLeuValAlaAlaTyrLeu 146
QY 778 ATAAAGAGATGACATGCTTTAAGATGAGCTTACAGATTGGAAGAAAGAACCT 837
Db 147 MetLysGluAsnAsnThrPheThrGluAlaAlaLeuSerIleLeuGlnLysArgSer 166
QY 838 ACTATATCTCCAACTTCAATTTTCTGGGCAACTCGCTGAGTAT 882
Db 167 GlyIleSerProAsnAlaAsnPheLeuArgLysLeuValTyr 181
RESULT 13
T16056
hypothetical protein F13D11.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001
C/Accession: T16056
R:Pulson, L.
submitted to the EMBL Data Library, November 1995
A/Description: The sequence of C. elegans cosmid F13D11.
A/Reference number: S69020
A/Accession: T16056
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-186 <FUL>
A/Cross-references: EMBL:U40939; NID:G1072175; PID:G1072176; PIDN:AAA81700.1; CESP:F13D11
C:Gene: CESP:F13D11.3
A:Gene: CESP:F13D11.3
A:Introns: 30/3; 57/3; 85/3; 125/3; 172/3
Alignment Scores:
Pred. No.: 9.46e-12 Length: 186
Score: 222.50 Matches: 51
Percent Similarity: 58.16% Conservative: 31
Best Local Similarity: 36.17% Mismatches: 50
Query Match: 13.68% Indels: 9
DB: 2 Gaps: 3
US-10-029-345A-108_COPY_538_1443 (1-906) x T16056 (1-186)
QY 478 ACCGGAATTCCTCCCAATCTTATCTT-----GGCTGCAGAGAGATGCTCAAC 528

Db 12 ThrGlnValArgProHisLeuPheLeuAlaGlyTyrCys-----IleThr 27
QY 529 AAGAGCTGATACAGAGAAATGGATGGTATGTATGTTAAATCCAGCTATACCTGTCCA 588
Db 28 ProSerLeuLeuLysGlnIlePheIleGlyValAlaAspCysThrIleLeuTyrThr 47
QY 589 AAGCTGACTTTATCCCGAGTCACTTCCGCGTGGCTGGCTGGATGACACTTTTGT 648
Db 48 LysPro-----IleLysGlyLeuAspArgIleGlyValAlaProValAspAsnThrLeu 65
QY 649 GAGAAATTTTGGCGGTGGACAAATGACATGATGATTTCTATGAGAAAGAGCTCC 708
Db 66 AlaLysIleThrGlnTyrRheGluProValValLysTyrIleGluAspAlaLysGln 85
QY 709 AATGATGTGTTCTAGTCACTGATGATGCTGGATGCTCCGCTCCGACCATGCTATC 768
Db 86 GlnHisAsnThrValIleTyrCysValAlaGlyValSerArgSerAlaThrLeuThrIle 105
QY 769 GCTTACATGATGAAGAGATGACATGTCTTAAATGAAAGCTTACAGATTGTGAAAGA 828
Db 106 ValTyrLeuMetValThrGluAsnLeuSerLeuGluGlnAlaTyrLeuGlnValAsnGln 125
QY 829 AAAAGACTACTTATCTCCAACTTCAATTTTCTGGGCCCACTCTGAGCTATGAGAG 888
Db 126 ValArgProIleIleSerProAsnIleGlyPheTrpArgGlnMetIleAspPheGluLys 145
QY 889 AAG 891
Db 146 Gln 146
RESULT 14
T21380
hypothetical protein F26A3.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T21380
R:McMurry, A.
submitted to the EMBL Data Library, August 1996
A/Reference number: Z19415
A/Accession: T21380
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-226 <WIL>
A/Cross-references: EMBL:Z78419; PIDN:CAB01700.1; GSPDB:GN00019; CESP:F26A3.4
A:Experimental source: clone F26A3
C:Gene: CESP:F26A3.4
A:Gene: CESP:F26A3.4
A:Map position: 1
A:Introns: 117/2; 150/3; 186/3
Alignment Scores:
Pred. No.: 7.48e-11 Length: 226
Score: 212.50 Matches: 47
Percent Similarity: 58.57% Conservative: 35
Best Local Similarity: 33.57% Mismatches: 57
Query Match: 13.06% Indels: 1
DB: 2 Gaps: 1
US-10-029-345A-108_COPY_538_1443 (1-906) x T21380 (1-226)
QY 478 ACCGGAATTCCTCCCAATCTTATCTTGCTGCCAGGAGATGCTCAACAGAGCTG 537
Db 15 SerGluIleValProGluLeuPheIle---CysGlyValSerAlaLeuSerLysAspGlu 33
QY 538 ATACAGCAGAAATGGGATGGTATGTATGTTAAATGCCAGCTATACCTGTCCAAAGCTGAC 597
Db 34 MetLysLysHisLysIleThrHisIleIleAsnAlaThrThrGluValProAsnLeuArg 53
QY 598 TTATCCCGGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAGAAATT 657
Db 54 SerLeuGlyAspIleGlnArgThrLysLeuTrpLeuGluAspArgPheProGlnThrTyrIle 73

```

QY      658  TTGCCGCGTGGACAAATCAGATATTCATTGAGGAAAGAAAAGCTCCATGGATGT  717
Db      74  TyrProHisIeuGluIueuGlnserXpGlnIleGlnAlaLeuIleAlaArgIlyGly  93
QY      718  GTTCTAATGACATCGTTTAAGCTGAGATCTCCGCTCCGACACAGCTATGCGCTCAATC  777
Db      94  ValIeuValHisCysValAlaClyAlaIserXpGlnIleSerIleCysIleuAlaPheLeu  113
QY      778  ATGAAAGGAGATGCAKATGCTTTAGATAGAAGCTTACAGATTTGTGAAAGAAAAGAAAGACT  837
Db      114  LeuIeuTyrArgCysArgIleuLeuArgIlnAlaTyrHisIleuMetIysSerIlyAspSer  133
QY      838  ACTAATATCCAAAACCTCAATTTTCTGGGCGCAACTCTCGATATGAGAGAAGAATTAAG  897
Db      134  MetValArgProAsnIeuGlyPheIerTyrArgIleuIleuAlaTyrIleuIleuAsnValIys  153

```

```
QY      874 CTGACCTATGAGGAAGATTAAg 897
      ::||| ||||| ||||| :::::
Db      142 ValMpyrGlnYsArgLeuArg 149

Search completed: June 21, 2004, 12:41:55
Job time : 21.7715 secs
```

C:\pep\c04f12\c04f12.8 - Caenorhabditis elegans
C:\pep\c04f12\c04f12.8 - Caenorhabditis elegans
C:\date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:\accession: T18915
R:\loyd, C.
submitted to the EMBL Data Library, November 1996
A:\reference number: Z19044
A:\accession: T18915
A:\status: preliminary; translated from GB/EMBL/DBJ
A:\molecule type: DNA
A:\residues: 1-272 <NRL>
A:\cross-references: EMBL:Z81461; PIDD:CAM03837.1; GSPDB:GM00019; CESP:C04F12.8
C:\experimental source: clone C04F12
C:\genetics:
A:\gene: CESP:C04F12.8
A:\map position: 1
A:\inlions: 53/1; 84/3; 204/2

Alignment Scores:	
Pred. No.:	1.9e-10
Score:	208.00
Percent Similarity:	54.05%
Best Local Similarity:	33.11%
Query Match:	12.78%
DB:	2
Length:	2722
Matches:	49
Conservative:	31
Mismatches:	62
Indels:	6
Gaps:	3

2025-108_COPY_538_1443 (1-906) X T18915 (1-272)

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:21; Search time 9.52534 Seconds
(without alignments)
9905.280 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 1627

Sequence: 1 atgagccacagatgatgctg9.....agaagataagaaccagact 906

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xld
-O=/cgn2.1/USPTO.spool.p/US10029345/runat.21062004.122815.4177/bpp.query.fasta.1.3278
-DB=SwissProt.42 -QFMT=fastan -SUFFIX=exp -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USR=US10029345.QCGN.1.1.33@runat.21062004.122815.4177 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: SwissProt.42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1552	95.4	665	DUSG_HUMAN	Q9by84 homo sapien
2	988	60.7	625	DUS8_HUMAN	Q13202 homo sapien
3	985	60.5	663	DUS8_MOUSE	O09112 mus musculu
4	469	28.8	482	DUSA_HUMAN	Q9y6w6 homo sapien
5	468	28.8	483	DUSA_MOUSE	Q9es80 mus musculu
6	467.5	28.7	381	DUS6_HUMAN	Q16828 homo sapien
7	462.5	28.4	381	DUS6_MOUSE	Q9dbb1 mus musculu
8	461.5	28.4	381	DUS6_RAT	Q64346 rattus norv
9	450	27.7	619	VHPI_CABEL	Q10038 caenorhabd
10	431	26.5	375	DUS4_CHICK	Q9w711 gallus gall
11	427	26.2	394	DUS4_HUMAN	Q13115 homo sapien
12	425	26.1	367	DUS1_MOUSE	P28563 mus musculu
13	423.5	26.0	395	DUS1_RAT	Q62767 rattus norv
14	418	25.7	314	DUS2_HUMAN	Q05993 homo sapien
15	415	25.5	367	DUS1_HUMAN	P28562 rattus norv
16	415	25.5	367	DUS1_RAT	Q64623 rattus norv
17	409	25.1	318	DUS2_MOUSE	Q05922 mus musculu
18	408	25.1	384	DUS9_HUMAN	Q99956 homo sapien

19	385.5	23.7	320	DUS7_HUMAN	Q16829 homo sapien
20	382.5	23.5	320	DUS7_MOUSE	Q91246 mus musculu
21	358.5	22.0	384	DUS7_RAT	O54838 rattus norv
22	357.5	22.0	280	DUS7_RAT	O63340 rattus norv
23	351	21.6	384	DUS5_HUMAN	Q16690 homo sapien
24	262	16.1	198	DUS5_HUMAN	Q95147 homo sapien
25	261	16.0	198	DUS5_MOUSE	Q9j147 mus musculu
26	205.5	12.6	188	DUS1_HUMAN	O8ne10 homo sapien
27	205.5	12.6	217	DUS1_HUMAN	O8wt22 homo sapien
28	205	12.6	223	STYX_MOUSE	O60969 mus musculu
29	204	12.5	223	STYX_HUMAN	O8wu10 homo sapien
30	202	12.4	364	PMV1_YEAST	Q02256 saccharomyc
31	201.5	12.4	313	PMV1_HUMAN	Q9y618 homo sapien
32	195.5	12.0	185	DUS3_HUMAN	P51452 homo sapien
33	194	11.9	185	DUS3_MOUSE	Q9d743 mus musculu
34	185	11.4	807	YB9T_YEAST	P38148 saccharomyc
35	183.5	11.3	276	PTP3_CHLEU	Q39451 chlamydomon
36	181.5	11.2	278	PMPI_SCHPO	O13453 chlamydomon
37	178.5	11.0	295	DUSF_HUMAN	O9h122 homo sapien
38	177	10.9	125	DUSF_MOUSE	O84422 mus musculu
39	171.5	10.5	339	DUSC_MOUSE	Q9d002 mus musculu
40	171	10.5	349	MSG5_YEAST	P38580 saccharomyc
41	158.5	9.7	340	DUSC_HUMAN	O9un16 homo sapien
42	157	9.6	209	YIL3_YEAST	P40479 saccharomyc
43	155	9.5	198	DUSD_HUMAN	Q9u116 homo sapien
44	152.5	9.4	597	PTPX_CANAL	P43078 candida alb
45	152	9.3	292	Y042_CABEL	P34680 caenorhabd

ALIGNMENTS

RESULT 1
DUSG_HUMAN
ID DUSG_HUMAN STANDARD; PRT; 665 AA.
AC Q9by84; Q9COG3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein kinase phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase
DE phosphatase 7) (MKP-7).
DE DUSP16 OR MKP7 OR KIAA1700.
OS Homo sapiens (human).
GN DUSP16
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21486429; Pubmed=11489891;
RA Nagata K., Shima H., Watanabe M., Kikuchi K.;
RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,
RT functions as a shuttle protein.";
RL J. Biol. Chem. 276:39002-39011 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21082932; Pubmed=11214970;
RA Nagata T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:347-355 (2000).
CC - FUNCTION: Involved in the inactivation of MAP kinases.
CC - CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC - CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC - SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC - SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC - SIMILARITY: Non-receptor class dual specificity subfamily.
CC - SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

[illegible]

CC -1- SIMILARITY: Contains 1 rhodanese domain.

CC -----
 CC The SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: X85518; CA64772.1; .
 DR HSP: Q16828; MKP.
 DR MGD: MGI:106626; Dusp8.
 DR InterPro: IPR000340; D8_phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR Pfam: PF00782; Dspc; 1.
 DR PRINTS: PR01764; MARKPHATASE.
 DR SMART: SM00195; Dspc; 1.
 DR SMART: SM00450; RHOD: 1.
 DR PROSITE: PS50206; RHODANES_3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase; Nuclear protein.
 FT DOMAIN 23 138 RHODANES.
 FT DOMAIN 162 432 PROTEIN-TYROSINE PHOSPHATASE.
 FT DOMAIN 452 459 POLY-ARG.
 FT DOMAIN 555 558 POLY-SER.
 FT DOMAIN 559 576 POLY-GLY.
 FT DOMAIN 577 600 POLY-SER.
 FT DOMAIN 311 552 PRO-RICH.
 FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
 SO SEQUENCE 663 AA; 68847 MW; 416F429A12C1FA7C CRC64;

Alignment Scores:
 Pred. No.: 5.03e-84 Length: 663
 Score: 985.00 Matches: 188
 Percent Similarity: 80.73% Conservative: 55
 Best Local Similarity: 62.46% Mismatches: 56
 Query Match: 60.54% Indels: 2
 Gaps: 2

US-10-029-345a-108_COPY_538_1443 (1-906) x DUS8_MOUSE (1-663)

QY 1 ATGGCCCATGAGTGAATGAGTCAATGTTT--ACTGAGAGGTTGGCTGCTGCTG
 Db 1 MetAlGlyAspArgLeuProArgLysValMetAspAlaLysLysLeuLeu 20
 QY 58 GAAAGTGAACGGAAGAAAGTCTGCTAATGATGATGCGCGCATTTTGGAAATCAATCA
 Db 21 ArgGlyGlyProGlyGlyProLeuValLeuAspSerLysSerPheValGluTyrAsnSer 40
 QY 118 TCCCATTTTGGAGCCATTAATCACTGCTCAAGCTTGAAGGAGGAGTTGCA 117
 Db 41 CysHsValLeuSerSerValAsnLeuLysCysSerLysLeuValLysArgLysLeu 60
 QY 178 GAGGCAAGTGTATTAATTAACAGAGCTATCAGATTCAGCGAAATAGAGTTGACAT 237
 Db 61 GlnGlyLysValThrIleAlaGluLeuIleGlnProLysThrArgSerLysValAspAla 80
 QY 238 GATTGACAGTGAAGGTTGATTAATGATCAAGTCCCAAGAGTGGCTGCTGCT 297
 Db 81 ThrGluProGlnAspValValValTyrAspGlnSerThrArgAspAlaSerValLeuAla 100
 QY 298 TCAACATGTTTCTCACTGATTAATTAATGAGTGAAGGAGGAGTTGACATGTTGCT 357
 Db 101 AlaAspSerPheLeuSerLysLeuLeuSerLysLeuAspGlyCysPheAspSerValAla 120
 QY 358 CTGCTGCAAGTGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417

Db 121 IleLeuThrGlyGlyPheAlaThrPheSerSerCysPheProGlyLeuGlyGlyLys 140
 QY 418 TCCACTTACTAGTCCCTACCC--TGCAATTTCTCAGCTTCTTACCTGTTGCCAATATGGG 474
 Db 141 ProLThrLeuProSerMetSerLeuSerGlnProCysLeuProValProSerValGly 160
 QY 475 CCAACCCGAAATTTCTCCCAATCTTATCTGTCGCCAGGAGATGCTTCAACAGAG 534
 Db 161 LeuThrArgIleLeuProHisLeuTyrLeuLysSerLysValLeuLeuLysAsp 180
 QY 535 CTGATACAGCAAGAAATGCGATTTGATGTGTTAATCCAGACTATTCCTGCAAGAGCT 594
 Db 181 LeuMetThrGlnAsnGlyLysSerTyrValLeuAsnAlaSerLysSerCysProLysPro 200
 QY 595 GACTTATATCCCGAGTCCATTTCTGCGCGTGCCTGTGATGACGCTTTTGTGAGAA 654
 Db 201 AspPheIleCysGlnSerArgPheMetArgLysLeuProIleAsnAspAsnTyrCysGlyLys 220
 QY 655 ATTTGCGCGTGTGAGCAATGATGATGATTTGATGAGAAAGCAAGCCCAATGGA 714
 Db 221 LeuLeuProThrLeuLeuLysSerLysLeuGluPheIleAspLysAlaLysLeuSerSerCys 240
 QY 715 TGTGTTTGAATGCACTGTTTACGTTGAGTCTCCGCTCCGACCATGCTATGCGCTAC 774
 Db 241 GlnValIleValHisCysLeuAlaGlyLysSerArgSerAlaThrIleAlaIleAlaTyr 260
 QY 775 ATCATGAAAGAGATGAGACATGCTCTTGAATGAGAGCTTACAGATTTGAGAAAGAAAGA 834
 Db 261 IleMetLysThrMetCysLysSerSerAspAlaTyrArgPheValLysAspArgLys 280
 QY 835 CCTACTAATATCTTCAACTTCAATTTTGTGAGGCAATCTCGACTGACTTGAAGAAAGAT 894
 Db 281 ProSerLysSerProAsnPheAsnPheLeuGlyGlnLeuLeuGlyTyrGluArgSerLeu 300
 QY 895 AAG 897
 Db 301 Lys 301

RESULT 4
 ID DUSA_HUMAN
 AC 0916W6; STANDARD; PRT; 482 AA.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase
 OS Dusp10 OR MKP5.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OK NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99321929; PubMed=10391943;
 RA Tanoue T., Moriguchi T., Nishida E.;
 RT "Molecular cloning and characterization of a novel dual specificity
 RU J. Biol. Chem. 274:19949-19956 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20065165; PubMed=10597297;
 RA Theodosiou A., Smith A., Gillieron C., Arkinstall S., Ashworth A.;
 RT "MKP5, a new member of the MAP kinase phosphatase family, which
 RL selectively dephosphorylates stress-activated kinases.";
 RN Oncogene 18:6981-6988 (1999).
 RP SEQUENCE FROM N.A.
 RX TISSUE=Brain, Lung, and Testis;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strauberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schutler G.D.,

RA	Ahtechul S.F., Jordan B., Butow K.H., Schaefer C.F., Bat N.K.,
RA	Hopkins R.F., Zeeberg H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Mansura K., Farmer M.A., Rubin G.M., Hong L.,
RA	Stratton M., Soares M.B., Bonaldi M.F., Casavant T.P., Scheetz T.E.,
RA	Brownstein M.J., Ueda T.B., Toshitsuki S., Carninci P., Frange C.,
RA	Rosa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA	Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA	Vallalon D.K., Worley D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahney J., Helton E., Kottelman M., Madan A., Rodriguez S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Buterfeitz J., S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA	Scherech A., Schein J.E., Jones S.J.M., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length
RT	human and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC	-1- FUNCTION: Involved in the inactivation of MAP kinases. Has a
CC	specificity for the MAPK1/MAPK2/MAPK3/MAPK4 subfamily.
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC	tyrosine + phosphate.
CC	-1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC	phosphate.
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC	-1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC	-1- SIMILARITY: Contains dual specificity subfamily.
CC	-1- SIMILARITY: Contains 1 rhodanese domain.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to license@sib-sib.ch).
CC	-----
DR	EMBL; AB026436; BAB1668.1; -
DR	EMBL; AF179212; AAD51857.1; -
DR	EMBL; BC031405; AAH31405.1; -
DR	HSSP; Q16828; IMKP.
DR	Genew; HGNC:1065; DUSP10.
DR	GO; GO:0005737; C:cytoplasm; TAS.
DR	GO; GO:0005634; C:nucleus; TAS.
DR	GO; GO:0004721; F:protein phosphatase activity; TAS.
DR	GO; GO:0007243; P:JNK cascade; TAS.
DR	GO; GO:0006740; P:protein amino acid dephosphorylation; TAS.
DR	GO; GO:0006950; P:response to stress; TAS.
DR	InterPro; IPR000340; DS phosphatase.
DR	InterPro; IPR008343; MAPK_phosph.
DR	InterPro; IPR001763; Rhodanese-like.
DR	InterPro; IPR000387; TYR_phosphatase.
DR	Pfam; PF00782; DSPC; 1.
DR	Pfam; PF00581; Rhodanese; 1.
DR	PRINTS; PRO1764; MARKPHPTASE.
DR	SMART; SMO0195; DSPC; 1.
DR	SMART; SMO0450; RHOD; 1.
DR	PROSITE; PS50206; RHODANESE_3; 1.
DR	PROSITE; PS50383; TYR_PHOSPHATASE_1; 1.
DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR	PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM	Hydrolyase; Nuclear protein.
FT	DOMAIN 168 285
FT	DOMAIN 384 453
FT	ACT_SITE 408 408
FT	SEQUENCE 482 AA; 52642 MW; A8CB74ABF9498CD4 CRC64;
SO	SEQUENCE 482 AA; 52642 MW; A8CB74ABF9498CD4 CRC64;

Query Match:	28.83%	Indels:	30
DB:	1	Gaps:	7
US-10-029-345A-108_COPY_538_1443 (1-906) x DUSA_HUMAN (1-482)			
OY	CTGCTAATTGATGACCCGCCCATTTGTGGAATACATATCATCCACATTTTGGAGCC		
Db	173 ValIleLeuAerCyArGrProPheMetGluTyAenLySerHisIleIleGlnLyAla		
OY	139 AATATCAACGTCGC---AAGCTTAATGGAGCGAAGGTTCACACAGACAAAGATGTA		
Db	193 HisIleAsnCyAlaMetLySileSerAaGhAghLeGlnGlnGlyLysIleThr		
OY	196 ACAGAGCTCATCCAG---CATTCAGGAAACATAAGTTGACATGATTGTCAGTCA		
Db	213 LeuAspLeuIleSerCyArGrGlnGlyLysAerSerPheLyArgIlePheSerLyS		
OY	253 GTTGATGATTACATCAAAAGCTCCCAAGATGTGGCTCTCTCTCTTCACAGCTGTTT		
Db	233 IleIleValLyArSpIleAsnThrAenGlnPProSerAaGhValMetProSerGlnPro		
OY	313 ACTGTAATCTGGGTAACTGAGGAAGACCTCAACTGTTCACCTGTTCAGCTG		
Db	253 HisIleValleuGlnSerLeuLyArGrGlnGlyLysGlnProLeuValLeuLySerLy		
OY	373 TTGCTGAGTCTCTCGTGTGTTTCCCTGCGCTCTGTGAAGAAA-----		
Db	273 LeuSerSerPheLySerIleAsnHisIleGlnAenLyCyArSpAerSerLeuGlnLeuGln		
OY	418 -----TCACGCTATGCGCTTACCTGACATTT		
Db	293 CyArGrGlnLyAlaLyGlnLyAlaSerAlaAlaSerSerLeuProGlnProIle		
OY	445 CAGCGTTCCTTACCT---GTTCGCAACATTTGGGCAACCCGAAATTTCCCAATCTT		
Db	312 ---ProThrThrProAerIleGlnAenAlaGlnLeuThrProIleLeuProPheLeu		
OY	502 CTGGCTGCTCCAGGAATGTCTTCAACAAAGACGTGATACAGACAAATGGGATTTGTT		
Db	331 LeuGlnLyAsnIleGlnIleAsnAlaGlnIleAsnLeuAerThrMetGlnAlyLeuAsnIleGly		
OY	562 GTGTTAATGTCGACG-----TATACCTGTCGCAACCTGACCTTTATCCG		
Db	351 ValIleAsnValThrThrHisIleuProLeuTyGHisTyGlnLyGlnGlyLeuPhe-----		
OY	607 GAGCTTCATTTCTCGTGTGCTGCTGTGAATGACAGCTTTGTGAGAAAATTTGCGCG		
Db	369 -----AsnTyLyArGrLeuProAlaThrAerSerLyAsnLyGlnIleAsnLeuArgIle		
OY	667 TTGGACAAATCACTAATTTCTATTTGAGGAAGAAAAGCTCCAAATGAGATGGTTTAC		
Db	387 PheGlnIleuAlaPheGlnPheIleGlnGlnAlaHisIleGlnCyGlyLySerLyLeuLeu		
OY	727 CACTGTTAGCTGGATCTCCGCTCCGACACATCGATATGCGCTATCATGAGAG		
Db	407 HisCyGlnHisIleLyAlaSerAerSerAlaThrIleValIleAlaTyLeuMetLyS		
OY	787 ATGCACATGCTTTAAGATGAGCTTACAGATTTTGGAAGAAAAGACCTACTATAT		
Db	427 ThrArgMetThrMetThrAspAlaTyLyPheValLySerLyLyAlaArgProIleIle		
OY	847 CCAAATCTCAATTTTTCGGGCGCAATCTCCGACATGAGAGAAATGATTAAGAAC		
Db	447 ProHisLeuAsnHisMetGlyGlnLeuLeuGlnPheGlnGlnAlaPheLeuAsnHis		
RESULT 5			
DUSA_MOUSE	STANDARD;	PRT;	483 AA.
AC	Q9ES60; Q9CZ9;		
AT	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DE	10-OCT-2003 (Rel. 42, Last annotation update)		
	Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.166)		

138
192
195
212
252
232
312
252
372
272
417
292
444
311
501
330
551
350
606
368
666
386
726
406
786
426
846
446

DB 408 HieCgclnlaaglyValserArgSerAlaThrIleValIleAlaTyrLeuMetlySHis 447
 QY 787 ATGACATCTCTTAACTGACCTTACAGATTGTGAAAGAAAAGACTACTATTC 846
 DB 428 ThArMetTherMetThArPaLaTyrLySpheValySglYsArgProIleIleSer 447
 QY 847 CCAACTGCAATTTCTGGGGCCAACTGCTGACACTGACATGAGAAAGATTAGAAC 900
 DB 448 ProshLeuAnPheMetGlyGlnLeuGluPheGluIleuAlaPheLeuAnPhe 465
 RESULT 6
 DUS6_HUMAN STANDARD; PRT; 381 AA.
 ID DUS6_HUMAN
 AC 016826; 075109; Q9BSH6;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
 DE phosphatase 3) (MKP-3) (dual specificity protein phosphatase PYST1).
 GN DUSP6 OR MKP3 OR PYST1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Forebrain;
 RX MEDLINE=96312959; PubMed=8670865;
 RA Groom L.A., Sneddon A.A., Alessi D.R., Dowd S., Keyse S.M.;
 RT "Differential regulation of the MAP, SAP and RK/p38 kinases by PyST1,
 RT a novel cytosolic dual-specificity phosphatase.";
 RL EMBO J. 15:3621-3632(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Liver;
 RX MEDLINE=99077745; PubMed=9858808;
 RA Furukawa T., Yatsuka T., Yousef E.M., Abe T., Yokoyama T.,
 RA Fukushima S., Soeda E., Hoshi M., Hayashi Y., Sunamura M., Kobari M.,
 RA Horii A.;
 RT "Genomic analysis of DUSP6, a dual specificity MAP kinase phosphatase,
 RT in pancreatic cancer.";
 RL Cytogenet. Cell Genet. 82:156-159(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT VAL-114.
 RC TISSUE=Colon, Kidney, Skin, and Stomach;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,
 RA Klausner R.D., Collins F.S., Moore T., Max S.L., Wang J., Heien F.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Heien F.,
 RA Diatchenko L., Marulisa K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bonak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Heien E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchan J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywiński M.I., Skalska U., Smilins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 204-347.
 RX MEDLINE=99140299; PubMed=10048930;
 RA Stewart A.E., Dowd S., Keyse S.M., McDonald N.O.;
 RT "Crystal structure of the MAPK phosphatase PyST1 catalytic domain and
 RT implications for regulated activation.";

RL Nat. Struct. Biol. 6:174-181(1999).
 CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
 CC family.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: 1.
 CC Event=Alternative splicing, Named isoforms=2;
 CC Name=1;
 CC IsoId=Q16828-1; Sequence=Displayed;
 CC Name=2; Synonyms=DUSP6-ALT;
 CC IsoId=Q16828-2; Sequence=VSP_005137;
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, X93920; CAA63813.1; -;
 DR EMBL, AB013601; BAA31968.1; -;
 DR EMBL, AB013383; BAA31968.1; JOINED.
 DR EMBL, AB013600; BAA31968.1; JOINED.
 DR EMBL, AB013382; BAA34369.1; -;
 DR EMBL, AB013602; BAA31969.1; -;
 DR EMBL, BC003143; AAH03143.1; -;
 DR EMBL, BC003562; AAH03562.1; -;
 DR EMBL, BC005047; AAH05047.1; -;
 DR EMBL, BC037236; AAH37236.1; -;
 DR PDB, 1MKP; 22-JUL-99.
 DR PDB, 1HZM; 25-JAN-02.
 DR GeneW; HGNC:3072; DUSP6.
 DR MIM; 602748; -;
 DR GO; GO:0005737; Cytoplasm; IDA.
 DR GO; GO:0005625; C:soluble fraction; TAS.
 DR GO; GO:0004722; Protein serine/threonine phosphatase activity; IDA.
 DR GO; GO:0004725; Protein tyrosine phosphatase activity; IDA.
 DR GO; GO:000188; P:inactivation of MAPK; IDA.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
 DR InterPro; IPR000340; DS phosphatase.
 DR InterPro; IPR008343; MAPK phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; Tyr phosphatase.
 DR Pfam; PF00782; DSpC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR SMART; SM00195; DSpC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANSE_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR HydroLase; Polymorphism; Alternative splicing; 3D-structure.
 FT DOMAIN 30 148 RHODANSE.
 FT 206 361 PROTEIN-TYROSINE PHOSPHATASE.
 FT 293 293 PHOSPHOCYSTEINE INTERMEDIATE.
 FT ACT SITE 134 279 Missing (in isoform 2).
 FT VARSPPLIC 134 279 /FTId=VSP_005137.
 FT VARIANT 114 114 L -> V.
 FT /FTId=VAR_015113.
 FT STRAND 208 211
 FT TURN 212 213
 FT STRAND 214 217
 FT TURN 219 220
 FT HELIX 225 230
 FT TURN 231 232

```

FT STRAND 233 238
FT STRAND 246 250
FT TURN 251 252
FT STRAND 253 257
FT TURN 264 265
FT HELIX 269 271
FT HELIX 272 284
FT TURN 285 286
FT STRAND 288 292
FT HELIX 298 312
FT TURN 313 313
FT HELIX 316 326
FT TURN 328 329
FT TURN 335 336
FT HELIX 337 345
FT TURN 346 346
SQ SEQUENCE 381 AA; 42333 MM; 03BC12252CE73B26 CRC64;

Alignment Scores:
Pred. No.: 1,08e-35 Length: 381
Score: 467.50 Matches: 110
Percent Similarity: 53.56% Conservative: 63
Best Local Similarity: 34.06% Mismatches: 105
Query Match: 28.73% Indels: 45
DB: 1 Gaps: 7

```

US-10-029-345a-108_copy_538_1443 (1-906) x DUS6_HUMAN (1-381)

```

QY 55 CTGGAAGAGTGAACCGGAAAGTCTGCTTAATGATACCGCGCATTTGTGAATACAT 114
Db 27 LeuGluLeuGlyAsnLnuArgLeuLeuLeuMetAspCysArgProGlnGluLeuYrGlu 46
QY 115 ACATCCCATATTTGGAAGCCATTATATCAACTGCTCCAAAGCTTATGAAGCAAGTTG 174
Db 47 SerSerHisIleGluSerAlaIleAsnValAlaIleProGlyIleMetLeuArgLeu 66
QY 175 CAACAGACAAAGGTGTTATACAGAGCTATCCAGCATTCACGC--AAACATAGATT 231
Db 67 GlnLysGlyAsnLeuProValArgAlaLeuPheThrArgGlyGlnAspArgAspArgPhe 86
QY 232 GACATGATTCG--AGTCAGAAAGTTGTAGTTTACAGTCAAGCTCCCAAGAT--GTT 285
Db 87 ThrArgArgCysGlyThrAspThrValValLeuIleAspLysSerSerAspTrpAsn 106
QY 286 GCTCTCTCTCTTCAGCTGTTTTCACCTGACTTCTGGGTAACTGAGAAGAGCTTC 345
Db 107 GlnAsnThrGlyGlyGluSerLeuLeuGlyLeuLeuLysIleLeuLysAspGluGly 126
QY 346 AACTCTTCACTGCTGCTGAGGTGGTTCGAGAGTTCTCGTGTTCCTGCGCTC 405
Db 127 CysArgAlaPheThrLeuGluGlyGlyPheSerLysPheGlnAlaGluPheSerLeuHis 146
QY 406 TGTGAAGAAATTCCTAGTCCCTACCTGC--ATTTCAGCCTTGTACTGCTGTT 462
Db 147 CysGlu--ThrAsnLeuAspLysCysSerSerSerSerProLeuProVal 164
QY 463 GCCAATCTGGG----- 474
Db 165 LeuGlyLeuGlyGlyLeuArgIleSerSerAspSerSerSerAspIleGluSerAspLeu 184
QY 474 ----- 474
Db 185 AspArgAspProAsnSerAlaThrAspSerAspLysSerProLeuSerAsnSerGlnPro 204
QY 475 -----CCAACCGCAATTTTCCCAATCTTTATCTTGGTGCAGGAGATGCTCTAC 528
Db 205 SerPheProValGluIleLeuProPheLeuIleYrLeuGlyCysAlaLysAspSerThrAsn 224
QY 529 AAGAGCTGATACAGCAGATGGATTGTTATGTTAAAGCCAGCTATACCTGTCCA 588
Db 225 LeuAspValLeuGluGlnPheGlyIleLysYrIleLeuAsnValThrProAsnLeuPro 244
QY 589 AAGCCT--CACTTATCCCGAGGCTCATTTCTGCGTGTGCTGTGAATGACAGCTTT 645

```

```

Db 245 AsnLeuPheGluAsnAlaGlyGluPheLysYrLysGlnIleProIleSerAspHisTrp 264
QY 646 TGTGAAGAAATTTTGGCGGTGGAGCAAAATTCATGATTTCAATGAGAAAGAACGCC 705
Db 265 SerGlnAsnLeuSerGlnPhePheProGluAlaIleSerPheIleAspGluAlaArgGly 284
QY 706 TCCATGATGATGTTCTGATGCTGCTGTTTACCTGGAGATCTCCGCTCCGACCATGCT 765
Db 285 LysAsnCysGlyValLeuValHisCysLeuAlaGlyIleSerArgSerValThrValThr 304
QY 766 ATCCGCTTACATCATGAAAGGATGAGATGCTGTTTGAATGAGTACGATTTGTGAA 825
Db 305 ValAlaIyrLeuMetGlnLysLeuAsnLeuSerMetAsnAspAlaIyrAspIleValLys 324
QY 826 GAAAAAAGACCTACTACATATCTCCAAATTTCTGAGGCAACTCCTGAGATATGAG 885
Db 325 MetLysLysSerAsnIleSerProAsnPheAsnPheMetGlyGlnLeuLeuAspPheGlu 344
QY 886 AAGAAAGATT 894
Db 345 ArgThrLeu 347

```

RESULT 7

```

ID DUS6_MOUSE STANDARD; PRT; 381 AA.
AC Q9DBB1; Q9D7L4;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
DE phosphatase 3) (MKP-3).
GN DUSP6 OR MKP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver, and Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Adachi S., Fukuda S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kaeukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirai L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima Y., Mazzarelli J., Monbets P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakomoto N.,
RA Sasaki H., Sato K., Schoenbach K., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschuler R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bat N.K.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.P., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Yoshitake S., Carninci P., Prange C.,
RA Rana S.S., Lottellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

```

```

RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunarathne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Wozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faley U., Helton E., Ketteman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywicki M.I., Skalska T., Smallus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
FL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
CC family (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to licenses@isb-sib.ch).
-----
DR EMBL, AK005062; BAB23786.1; -
DR EMBL, AK009131; BAB26093.1; -
DR EMBL, BC003869; AAH03869.1; -
DR HSSP, Q16828; IMKP.
DR MGD, MGI:1914853; Duap6.
DR InterPro, IPR000340; DS_phosphatase.
DR InterPro, IPR008343; MAPK_phosph.
DR InterPro, IPR001763; Rhodanese-like.
DR InterPro, IPR000387; TYR_phosphatase.
DR Pfam, PF00782; DSPc; 1.
DR PRINTS, PRO1764; MARKPHPTASE.
DR SMART, SMO0195; DSPc; 1.
DR SMART, SMO0450; RHOD; 1.
DR PROSITE, PS50206; RHODANESE_3; 1.
DR PROSITE, PS00183; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE, PSS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE, PSS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolyase.
FT DOMAIN 30 148 RHODANESE.
FT FT 206 381 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 293 293 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
FT CONFLICT 22 22 W->G (IN REF. 1; BAB26093).
FT CONFLICT 34 34 L->F (IN REF. 1; BAB26093).
SQ SEQUENCE 381 AA; 42407 MW; TEAIFB154FAD2DA CRC64;

```

Alignment Scores:

Pred. No.:	3,17e-35	Length:	381
Score:	462.50	Matches:	109
Percent Similarity:	53.56%	Conservative:	64
Best Local Similarity:	33.75%	Mismatches:	105
Query Match:	28.43%	Indels:	45
Gaps:	1	Gaps:	7

```

US-10-029-345A-108_COPY_538_1443 (1-906) x DUS6_MOUSE (1-381)
DY CTGGAAGTCGAAACCGAAAAAGTCCTGTATTATGATACCGGCATTTGGTAGTAACAAT 114
| | | | | : : : : : | | | | |
DB 27 LeuGlutIeuGliYlaengluAlargLeuleuMetAhpCysArgPrGlnGluIeuTryGlu 46
:::||||||| |||||:::|:::|
115 ACATGCCCATTTTGGAGCAAGCATTAATATCAACTGCTCCAAGCTTATGAGCAAGGTTG 174
:::||||||| |||||:::|:::|

```

[illegible]

```

Qy 55 CTGGAAAGTGGAAACGGAAAAAGCTGCTAAATTGATTCAGCGGACATTTTGGAATACAAAT 114
Db 27 LeuGluLeuGlyAsnGluGlnLeuLeuMetAspCyAsnProGlnGluLeuTyGlu 46
Qy 115 ACATCCCACTTTTGGAAAGCCTTAATATTCACATGCTTCCAAAGCTTATGAAAGCGAAAGTTG 174
Db 47 SerSerIleIleGluSerIleAlaIleAsnValAlaIleProGlyIleMetLeuArgLeu 66
Qy 175 CAAGAGGACAAAGCTGTAAATTATTCAGAGCTTCCAGCATCC--TCACGGAACAACTAAGTT 231
Db 67 GlnHisGlyAsnMetProValArgAlaLeuPheThrArgCysGluAspArgAspArgPhe 86
Qy 232 GAACATTGATTC--AGTCAGAAAGTTGGATTGATTCAGATCAAAAGCTCCCAAGAT--GTT 285
Db 87 ThrArgArgCysGlyThrAspThrValValLeuTyAspGluAsnSerSerAspTrpAsn 106
Qy 286 GCCTCTCTCTCTTCAGACCTGTTTCTACTGTACTTCTGGATTAACCTGAGAAAGAGCTTC 345
Db 107 GlnAsnThrGlyGlyGluSerValLeuGlyLeuLeuLeuTyGlyLeuTyAspGlnGly 126
Qy 346 AACTCTGTTCACCTGCTTCGAGGAGGGATTTGGCTGAGATTCTCGTGTGTTTCCCGGCTC 405
Db 127 CysArgAlaPheTyLeuGlnGlyGlyPheSerTyPheGlnAlaGluPheAlaLeuHis 146
Qy 406 TGTGAAGGAATAATCACTACTAGTCCCTACTCC--ATTTCAGCCTTCTTACCTGTT 462
Db 147 CysGlu-----ThrAsnLeuAspGlySerCysSerSerSerProLeuProVal 164
Qy 463 GCCAAACTTGGG----- 474
Db 165 LeuGlyLeuGlyLeuArgIleSerSerAspSerSerAspIleGluSerAspLeu 184
Qy 474 ----- 474
Db 185 AspArgAspProAsnSerAlaThrAspSerAspGlySerProLeuSerAsnSerGlnPro 204
Qy 475 -----CCAAACCGAATCTTCCCAACTTTATCTTGGCTGCGCAGCGAAGATGCTCTCAAC 528
Db 205 SerPheProValGlnIleLeuProPheLeuTyLeuGlyCysAlaIleAspSerThrAsn 224
Qy 529 AAGAGACGTATACACGAAATGGATGGATGGATTGATTGTTTAAATGACGACTATACCTGTCCA 588
Db 225 LeuAspValLeuGlnGluPheGlyIleTyTyIleLeuAsnValThrProAsnLeuPro 244
Qy 589 AAGCCT--GACTTTATCCCGAGTCTCACTTTCCTGCGGTGCTGCGAATGACAGCTT 645
Db 245 AsnLeuPheGluAlaGlnIleGlyLeuPheTyTyTyGlnIleProIleSerAspHisTrp 264
Qy 646 TGTGGAAGAAATTTGGCGGTGTTGACAAATCAGATGATTCATTTGAGAAAGCAAAAGCC 705
Db 265 SerGlnAsnLeuSerGlnPhePheProGlnAlaIleSerPheIleAspGluAlaAspGly 284
Qy 706 TCCATGATGATGTTTCTAGTCAGCTGTTTACGTGGATCTCCGCTCGCCACCATCGCT 765
Db 285 LysAsnGlyGlyAlaLeuValHisCysLeuAlaGlyIleSerArgSerValThrValThr 304
Qy 766 ATCGCCCTACATCATGAAAGAGATGACATGTCTTTAGATGAACCTTACGATTTTGAGAA 825
Db 305 ValAlaIleTyLeuMetGlnHisLeuAsnLeuSerSerAsnAspIleArgIleValHis 324
Qy 826 GAAAAAAGACTACTATATCTCCAAACTTCAATTTCTGGCGCAACTCTTGACATATGAG 885
Db 325 MetLysLysSerAsnIleSerProAsnPheAsnPheMetGlyGlnLeuLeuAspPheGlu 344
Qy 886 AAGAAGATT 894
Db 345 ArgThrLeu 347

```

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Protein-tyrosine phosphatase vhp-1 (EC 3.1.3.48).
 GN vhp-1 OR P08B1.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pseudocercariae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Chiasse S.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL/ U23178; AAC6719.1; -.
 DR PIR; T15969; T15969.
 DR HSSP; Q16828; IMKP.
 DR WormBep; F08B1.1; CE01899.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR SMART; SM00195; DSPC; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KM HydroLase.
 FT DOMAIN 200 269 PROTEIN-TYROSINE PHOSPHATASE.
 FT DOMAIN 92 95 POLY-SER.
 FT DOMAIN 351 354 POLY-SER.
 FT DOMAIN 465 472 POLY-SER.
 FT DOMAIN 483 488 POLY-SER.
 FT ACT SITE 224 224 PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT SIMILARITY).
 SQ SEQUENCE 619 AA; 66354 MW; 369B326F61D0529 CRC64;
 Alignment Scores:
 Pred. No.: 5.09e-34 Length: 619
 Score: 450.00 Matches: 90
 Percent Similarity: 69.66% Conservative: 34
 Best Local Similarity: 50.56% Mismatches: 48
 Query Match: 27.66% Indels: 6
 DB: 1 Gaps: 4
 US-10-029-345a-108_COPY_538_1443 (1-906) x VHP1_CABEL (1-619)
 QY 367 GGTGGTTGCTGAGTCTCTCGTGTTCCTGCGCTCTGTGAAGAAATCC----- 420
 DB 99 GYGLYPhelYseGlnhEnaGlnGlnYrProGlnLeuCySeGlnuSerGlnuGlyMet 118
 QY 421 ACTTACTGCTTACCTGATCTTCTGAGCTTGTGTTA-----CCTGTGCGCAACATTGGG 474
 DB 119 ThirArGlnProlInserLeuSerGlnProCyLeuSerGlnProThirGlyAsp--Gly 137
 QY 475 CCAACCGAATTCCTCCATCTTATCTTGGCTGCAGGAGGATGCTCCACAGAGAG 534
 DB 138 IIEThirLeuIIEThirProhEnIIErYrLeuGlySeIIErInIEhSerLeuAspGlnuTr 157
 QY 535 CTGATACAGCAGATGGATGGTATGTTAAATCCAGTATACCTGTCCAAAGCCT 594
 DB 158 MetLeuAspAlaLeuAspIIErSerValIIEhAsnLeuSerIEThirCySProlYSer 177

QY 595 GACTTATCCCGAG---TCATATTCGCGGTGCTGTGAATGACACTTTTGNAG 651
 DB 178 ValCylLeuIIEhSerGlnuAspLeuAspPheMetArgIIErProValAsnAspSerYrGlnGlu 197
 QY 652 AAAATTTTGGCGGTGGTGGCAATATGATGATTTCTATGAGAAAGCAAAAGCTCCCAT 711
 DB 198 IIEhSerProIIErTherProMetAlaIIErGlnuPheLeuGlnuIIErCySeAspAlaGly 217
 QY 712 GATATGTTCTAGTGCATGTTTAACTGGATCTCCGCTCCGACCATGCTATGCG 771
 DB 218 IIEhSerCylLeuIIEhSerCylLeuAlaGlyIIErSerArgSerProThirLeuAlaIIEr 237
 QY 772 TATATGATGAAGATGATGACATGCTTATGATGAAGATGATGATTTTGAAGAAAA 831
 DB 238 TyrlIEhMetArgIIErTherMetGlySerAspAlaIIErYrValIIErGlnuArg 257
 QY 832 AGACTATATATCTCCAAACTCAATTTTCTGGCGCACTCTGACATATGAG 885
 DB 258 ArgProSerIIErSerProhEnPhehEnPheMetGlyGlnLeuGlnuIIErGlnu 275
 RESULT 10
 DUS4 CHICK STANDARD; PRT; 375 AA.
 AC 09PW71;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
 DE phosphatase-2) (MKP-2).
 GN DUSP4 OR MKP2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White Leghorn;
 RX MEDLINE=20379359; PubMed=10918612;
 RA Fu S.-L., Waha A., Vogt P.K.;
 RT Identification and characterization of genes upregulated in cells
 RT transformed by v-jun.";
 RL Oncogene 19:3537-3545(2000).
 CC -1- FUNCTION: Regulates mitogenic signal transduction by
 CC dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
 CC and ERK2 (By similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF167296; AAD4656.1; -.
 DR HSSP; Q16828; IMKP.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.

[2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93360956; PubMed=8355678;
 RA Nouguchi T., Metz R., Chen L., Mattei M.-G., Carrasco D., Bravo R.;
 RT "Structure, mapping, and expression of erp, a growth factor-inducible
 RT gene encoding a nontransmembrane protein tyrosine phosphatase, and
 RT effect of ERP on cell growth.";
 RL Mol. Cell. Biol. 13:5195-5205(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Czech II; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L.D., Shenen C.M., Schler G.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshyuki S., Carninci P., Prange C.J.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abrahams R.D., Mullaly S.J.,
 RA Bosak S.A., McKernan P.U., McKernan K.J., Malek A.M., Gay L.J., Hulik S.W.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez R.A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko I., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.L., Skalska U., Smallus D.E.,
 RA Schenck A.V., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=94037096; PubMed=8221888;
 RA Sun H., Charles C.H., Lau L.F., Tonks N.K.;
 RT "MKP-1 (3CH134), an immediate early gene product, is a dual
 RT specificity phosphatase that dephosphorylates MAP kinase in vivo.";
 RL Cell 75:487-493(1993).
 CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
 CC kinase ERK2 on both Thr-183 and Tyr-185.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- INDUCTION: By growth factors.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X61940; CA443944.1; -;
 DR EMBL; S64851; AA827882.1; -;
 DR EMBL; BC006967; AA06967.1; -;
 DR PIR; A54681; S24411.
 DR HSP; Q16828; IMKP.
 DR MGD; MGI:105120; Dusp1.
 DR InterPro: IPR000340; DS phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR Pfam; PF00782; DSPc; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; NAKPDPHTASE.
 DR SMART; SM00195; DSPc; 1.

DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PSS0206; RHODANES_3; 1.
 DR PROSITE; PSS0383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase; Cell cycle.
 FT DOMAIN 20 137
 FT DOMAIN 175 367
 FT ACT SITE 258 258
 FT MOTIF 258 258
 FT MOTIF 258 258
 SQ SEQUENCE 367 AA, 39369 MW, 5085F90FEBBD19AB CRC64;
 Alignment Scores:
 Pred. No.: 1.01e-31 Length: 367
 Score: 425.00 Matches: 104
 Percent Similarity: 50.17% Conservative: 48
 Best Local Similarity: 34.32% IdMatches: 129
 Query Match: 26.12% Indels: 22
 DB: 1 Gaps: 5
 US-10-029-345A-108_COPY_538_1443 (1-906) x DUS1_MOUSE (1-367)
 QY 43 TTGGTGGCTCTGCTGGAAGTGAACGGAAGGCTGCTTAATTGATGCGGCATT 102
 Db 13 LeuArgAlaLeuLeuArgGluGluAlaValAlaGlnCysLeuLeuAspCysArgSerPhe 32
 QY 103 GTGAATACAAATACATCCACATTTTGGAGCCATTAAATACATGCTCCAGCTTATG 162
 Db 33 PheAlaPheAspAlaGlyHisIleAlaGlySerValAsnValArgPheSerThrIleVal 52
 QY 163 AAGCGAAGTTCACACAGACAAAGTGTAAATTCACAGCTCATCCAGCATTCAGCGAA 222
 Db 53 ArgArgArgAlaValGlyAlaMetGlyLeuGlnHisIleValProAsnAlaGluLeuArg 72
 QY 223 CATAGGTTCACATTGANTGACAGTCACAGAGTTGATTTACATGACAGTCCCAAGT 282
 Db 73 GlyArgLeuLeuAlaGlyAlaIleValAlaValAlaValLeuLeuAspGluArgSerIle 92
 QY 283 GTTGGCTCTCTCTTTCAGACTGTTTTCACAGTCTTCTGAGTCTTCTGAGTCTTCTG 336
 Db 93 LeuAspGlyAlaValArgAspGlyThrLeuAlaLeuAlaAlaGlyAlaLeuCysArgGlu 112
 QY 337 AAGAGCTTCACCTGTTGACCTGCTGAGCTGAGTGGCTTCTGAGTCTTCTGTTGTT 396
 Db 113 AlaArgSerThrGlnValPhePheLeuGlnGlyGlyArgValAlaPheSerAlaSerCys 132
 QY 397 CTTGGCCTCTGTCAGAGAAATTCACCTGATGCTTACCTGATTCATTTCTCAGGCT 450
 Db 133 ProGluLeuCysSerIleGlnSerThr-----ProThrGlyLeuSerLeuProLeuSer 150
 QY 451 -----TGCTTACCTGTTGCCAATTT----- 471
 Db 151 ThrSerValProAspSerAlaGluSerCysSerCysSerThrProLeuIleArg 170
 QY 472 -----GGGCGCAACCGCAATTTCTCCCAATCTTATCTGCTGCGCAGAGATGTCCT 525
 Db 171 GlnGlyGlyProValGluIleLeuSerPheLeuTyrLeuGlySerIleValIleValSerIleAsn 190
 QY 526 AACCAAGAGCTGATACAGCAGAAATGATGTTATGTTGTTAAATGCCAGTTATCTGT 585
 Db 191 ArgIleAspMetLeuAspAlaLeuGlyIleThrAlaLeuIleAsnValSerIleAsnCys 210
 QY 586 CCAAGAGCTGATTCATCCCGAGTCTCATTTCTGCGTCCGTCGTCGTAATGACAGCTTT 645
 Db 211 ProAsn---HisPheGluGlyHisIleValGlySerIleValGluAspAsnHis 229
 QY 646 TGTGAGAAATTTTGGCCGTGTTGACAAATCATGATGATTCATGAGAAAGCAAGCC 705
 Db 230 LysAlaAspIleSerSerTrpPheAsnGluAlaIleAspPheIleAspSerIleValAsp 249
 QY 706 TCCAAATGATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 765
 Db 250 AlaGlyGlyArgValPheValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCys 269


```

QY 766 ATGCGCTCATCATGAGAGATGACATGCTTTAGATGAAGCTTACAGATTTTGAAA 825
Db 270 LeuAlaTyrLeuMetCAGTThrAraArgValLysLeuMetArgLysLysValLys 289
QY 826 GAAAAAGACCTTACTATATCTCCAAACTTATTTCTGGAGCCAACTCTGATATGAG 885
Db 290 GlnArgArgSerLeuLeuSerProAraPheSerPheMetGlyGlnLeuLysGlnPheGlu 309
QY 886 AAGAGATT 894
Db 310 SerGlnVal 312

RESULT 13
DUS4_RAT
ID DUS4_RAT STANDARD; PRT; 395 AA.
AC 062767;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
DE phosphatase-2) (MKP-2).
DE DUSP4 OR MKP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Phenochromocytoma; PubMed=7782322;
RA MEDLINE=95301550;
RA Miera-Press A., Rlm C.S., Yao H., Robertson M.S., Stork P.J.S.;
RT "A novel mitogen-activated protein kinase phosphatase. Structure,
RT expression, and regulation."
RL J. Biol. Chem. 270:14587-14596(1995).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
CC and ERK2 (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed at moderate levels in nearly all
CC tissues and cells including brain, spleen, and testes with the
CC higher expression in the heart and lung and lower expression in
CC skeletal muscle and kidney. Undetectable in liver. Expressed in
CC many areas of the brain with very strong expression in the
CC hippocampus, piriform cortex, and the suprachiasmatic nucleus.
CC -1- INDUCTION: By mitogens and by stress.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U23438; AAC52493.1; -.
CC HSSP: Q16828; IMKP.
CC InterPro: IPR000340; DS_Posphatase.
CC InterPro: IPR008343; MAPK_phosph.
CC InterPro: IPR001763; Rhodanese-like.
CC InterPro: IPR000387; TYR_phosphatase.
CC Pfam: PF00782; DSpC; 1.
CC Pfam: PF00581; Rhodanese; 1.
CC PRINTS: PR01764; MAPKPHPTASE.
CC SMART: SM00195; DSpC; 1.

```

```

DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE 3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE 1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE 2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; Nuclear protein.
FT DOMAIN 42 160 RHODANSE.
FT DOMAIN 198 395 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 281 281 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 395 AA; 43187 MW; A90BFFD378A050FD CRC64;

Alignment Scores:
Pred. No.: 1 41e-31 Length: 395
Score: 423.50 Matches: 98
Percent Similarity: 52.30% Conservative: 61
Best Local Similarity: 32.24% Mismatches: 118
Query Match: 26.03% Indels: 27
DB: Gaps: 7

US-10-029-345a-108_COPY_538_1443 (1-906) x DUS4_RAT (1-395)
QY 46 GTGGCTCTGCTGGAAAGTGAAGCAAGAAAGTCTGTAATTGATGACCGGCATTGTG 105
Db 38 LeuGlyLeuLeuSerGlyGly-----LysCysLeuLeuLeuMetCysArgProPheLeu 55
QY 106 GAATACAAATACATCCACATTTTGGAAAGCATTAATATCACTGTCGCAAGCTTAAGAAG 165
Db 56 AlaHisSerAlaGlyTyrIleArgGlySerValAsnValArgCysAsnThrIleValArg 75
QY 166 CGAAGGTGACAGACAGCAAGAGTTAATTACAGAGCTGTC-----CAGCATTCAGCG 219
Db 76 ArgArg---AlaLysGlySerValSerLeuGlnIleLeuProAlaGluGluVal 94
QY 220 AAACATAGGTTGACATTCATTGTCAGTACAGAGTTGATGATTCAGATCAAGCTCCAA 279
Db 95 ArgAlaArgLeuArgSerGlyLeuTyrSerAlaValIleValTyrMetArgLysArgPro 114
QY 280 GATGTGCTCTGCTCTCTTCAAGCTGTTTTCAGTCTGACTTGTGGTAACCTGGAGAG 339
Db 115 ArgAlaGluSerLeuArgLysSerThrValSerLeuValAlaGlnAlaLeuArg 134
QY 340 AGCTTC-----AAGCTGTTCACCTGTTGACAGGAGGAGTTCGATGCTTCGTTG 393
Db 135 AsnAlaGluArgThrAspIleCysLeuLeuLysGlyGlyTyrGluArgPheSerGly 154
QY 394 TTCCCTGGGCTGTGTGAAGAAATCCAGTACGTCCCTACCTGCAATTTCTGCACTTGC 453
Db 155 TyrProGluPheCysSerIleThrIleValLeu-----AlaAlaIleProProVal 172
QY 454 TTACTGTGGCCAACTT----- 471
Db 173 ProProSerThrAsnGluSerLeuArgLeuGlyCysSerSerCysGlyThrProLeuIle 192
QY 472 -----GGGCAACCCGAATTCCTCCCAATCTTATCTGAGGCGCAAGCATGTC 522
Db 193 ArgGlnGlyIleProValGluLeuProPheLeuTyrLeuGlySerAlaTyrGlnVal 212
QY 523 CTCACAAAGACCTGATACAGAGAGATGGATGGTATGTTAATGACAGCTTATACC 582
Db 213 AlaArgArgAspMetCysMetArgLysGlyIleThrAlaLeuLeuMetAsnLysSerArg 232
QY 583 TGTCCAAAGCTGACTTATCCCGAGTCAATTCTCGCGGTGCGCTGTAATGACAGC 642
Db 233 CysProAsn---HisPheGluGlyIleTyrGlnTyrLysCysValIleProValGluAsn 251
QY 643 TTTTGGAGAAATTTTGCAGGTTGAGCAAAATGATGATTTTCAAGAAAGCAAA 702
Db 252 HisLysValAspIleSerSerThrPheMetGluAlaIleGluTyrIleMetArgValLys 271
QY 703 GCGTCCATGATGATGTTCTAGTGCATGTTTGAAGTGGATGCTCCGCTCGGCAATC 762
Db 272 AspCysArgGlyArgValLeuValHisCysGlnAlaGlyIleSerArgSerIleThrIle 291

```

```

QY 763 GCTATGCGCTTACATCATGAGATGCATGCTTATGATGAGCTTACAGATTGTC 822
Db 292 CysLeuAlaIyrLeuMetCysIysAlrValArgLeuGluGluAlaPheGluPheVal 311
QY 823 AANAGAAAAAGACCTACTATATCTCCAACTTCAATTTTCGGCCCACTCTGAGCTAT 882
Db 312 LysGlnArgArgSerLeuIleSerProAsnPheserPheMetGlyGlnLeuGlnPhe 331
QY 883 GAGAGAAGAAAT 894
Db 332 GluSerGlnVal 335

RESULT 14
DUS2_HUMAN
ID DUS2_HUMAN STANDARD; PRT; 314 AA.
AC 005923;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 2 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Dual specificity protein phosphatase PAC-1).
GN DUSP2 OR PAC1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=93206122; PubMed=7681221;
RA Rohan P., Davis P., Moskaluk C.A., Kearns M., Krutzsch H.,
RA Stebenlist U., Kelly K.,
RT "PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.",
RL Science 259:1763-1766 (1993).
RN 12;
RP SEQUENCE FROM N.A.
RX MEDLINE=96070437; PubMed=7590752;
RA Yi H., Morton C.C., Meremowicz S., McBride O.W., Kelly K.,
RT "Genomic organization and chromosomal localization of the DUSP2 gene,
RT encoding a MAP kinase phosphatase, to human 2p11.2-q11.",
RL Genomics 28:92-96 (1995).
RN 13;
RP SEQUENCE FROM N.A.
RC TISSUE=B-cell;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueteow K.H., Schaefer C.F., Bhat N.K.,
RA Diatchenko L., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.D., Hong L.,
RA Brownstein M., Usdin T.B., Toshiyuki S., Carrinci P., Prange T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Richards S., Morley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Gay L.D., Hulik S.W.,
RA Fahy J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalek U., Smilus D.E.,
RA Schermer A., Schein J.E., Jones S.J.W., Skalek U., Smilus D.E.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases
CC ERK1 and ERK2.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear.

```

```

CC -1- TISSUE SPECIFICITY: In hematopoietic tissues.
CC -1- INDUCTION: By mitogens.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J11329; AAA50779.1; -
DR EMBL; U23853; AAA86112.1; -
DR EMBL; BC007771; AAH07771.1; -
DR PIR; A57126; A57126.
DR PDB; 1IKZ; 30-MAY-02.
DR Genew; HGNC:3068; DUSP2.
DR MIM; 603068.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.
DR GO; GO:0008330; F:protein tyrosine/threonine phosphatase acti. . .; TAS.
DR GO; GO:0000188; P:inactivation of MAPK; TAS.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK-phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; RHOD; 1.
DR PROSITE; PS00450; RHOD; 1.
DR PROSITE; PS0206; RHODANES 3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00564; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00564; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase; Nuclear protein; 3D-structure.
FT DOMAIN 23 144 RHODANES.
FT ACT_SITE 237 302 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 257 257 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 314 AA; 34399 MW; FDD3543CDE10CA5 CRC64;
Alignment Scores:
Pred. No.: 4,43e-31
Score: 418.00 Length: 314
Percent Similarity: 50.83% Matches: 109
Best Local Similarity: 35.97% Conservative: 45
Query Match: 25.69% Mismatches: 105
Db: 1 Indels: 44
Gaps: 8
US-10-029-345a-108_COPY_538_1443 (1-906) x DUS2_HUMAN (1-314)
QY 70 GAAAAAGTCTCTTATGATGATGAGCGGCATTTGTGGAATACATACCCAGATTTC 129
Db 25 GluArgThrLeuLeuLeuAspCysArgProPheLeuAlaPheCysArgArgHsValArg 44
QY 130 GAAGCATTAATTAATCACTGCTCCAGACTTATGACGAGAGG----- 171
Db 45 AlaAlaArgProValProTPrPheAlaLeuLeuArgArgAlaArgGlyProProAla 64
QY 172 -----TTGACACAGACAAAGTTGTATTACAGAGCTCATCCAGATTCA 216
Db 65 AlaValLeuAlaCysLeuLeuProAspArgAlaLeuArgThrArgLeuValAlaArgGlyGlu 84
QY 217 GCGAACAATAGGTTGACATTGTCACAGAGGTTGTATTACAGCAAAAGCTCC 276
Db 85 Leu-----AlaArgAlaValAlaLeuAspGluGlySer 95
QY 277 CAAGATGTCCTCTCTCTCTTCAAGAC-----TCTTTCTCACTGACTTCTG 324

```


..._COPI_538_1443 (1-906) X DUS1_HUMAN (1-367)

12 507/
 43 TTGGTGGCTCTGCTGGAAAGTGGAAACGGAAAAATGCTCTCTATTTGATTAATGAGCCGGCCATT 102
 Db 13 LeuArgAlaLeuLeuGlyGluArgAlaAlaGlnCybLeuLeuLeuAspCybArgSerPhe 32
 103 GTGGAAATACAAATACATCCCAATTTTGGAAAGCCATTATATCAATGCTGCCAACCTTATG 162
 Oy 33 PheAlaPheAsnAlaGlyHisIleAlaGlySerValAsnValAlaArgPheSerThrIleVal 52
 Db 163 AAGGAAAGGTTTGCACAGGACCAAGCAAAAGTTTATTTACAGAGCTCATCCAGCTTACGGGAAA 222
 Oy 53 ArgArgArgArgAlaLeuGluAlaLeuMetGlyLeuGluHisIleValAlaProAsnAlaGluLeuArg 72
 Db 223 CATTAAGTTTGCACATTTGATTCGACGTACAGAAAGTTTGATTTACATGCATCAAGCTCCCAAGT 282
 Oy 73 GluArgLeuLeuAlaGluAlaArgHisValAlaValAlaLeuLeuAspGluArgSerAlaAla 92
 Db 283 GTTGCCCTCTCTCTCTTCCAGACTGTTTTCTCACTCTACTCTTGCGTAACTG-----GAG 336
 Oy 93 LeuAspGluAlaAluArgAspGluThrIleuAlaLeuAlaAlaGluAlaLeuCybArgGlu 112
 Db 337 AAGAGCTTCAACTCTGTTCACCTCTTGACGGAGGCTGGTTGCTGAGTCTCTCGTTGGTTTC 396
 Oy 113 AlaArgAlaAlaGlnValPhePheLeuLeuGlyGlyThrGluAlaPheSerAlaSerCyb 132
 Db 397 CTGGCCCTCTGGAAAGGAAA-----TTCACTCTA 426
 Oy 133 ProGluLeuCybSerLysGlnSerThrProMetGluLeuSerLeuProLeuSerThrSer 152
 Db 427 GTCCCTACCTGCATTTCTGCAGCTTGCTTACCTGATCCGTGGCAACAT----- 471
 Oy 153 ValProAspSerAlaGluSerGlyCybSerSerCybSerThrProLeuTyrAspGlnGly 172
 Db 472 GGGGCAACCCGGAATTTCTCCCAATCTTTATCTTGGCTCCGACGAGATGTTCTCAACAG 531
 Oy 173 GlyProValGluIleuProPheLeuTyrLeuGlySerAlaTyrHisAlaSerArgLys 192
 Db 532 GAGGTATACAGACAAATGGATGGTATGTGTTAATGAGCACTATACCTGTCCAAAG 591
 Oy 193 AspMetLeuAspAlaLeuGluYylLeuThrAlaLeuAlaSerAlaAsnCybProAsn 212
 Db 592 CTGACTTATATCCCGAGCTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTGAG 651
 Oy 213 ---HisPheGluGluHisTyrGlnTyrTyrLeuSerIleProValGluLysAsnHisAluAla 231
 Db 652 AAAATTTGCGCGTGTGGACAAATACATGATTTTATGGAAGAAGCAAAAGCCGCCAAT 711
 Oy 232 AspIleSerSerThrPheAsnGluAlaIleAspPheIleAspSerIleLysAsnAlaGly 251
 Db 712 GGATGTGTTTAGTGCACTGTTTAGCTGGAGGATCTCCCGCTCCGCCACCACTAGCTAGCC 771
 Oy 252 GluArgValAlaPheValHisCysGlnAlaGlyIleSerValSerAlaThrIleCysLeuAla 271
 Db 772 TACATCATGAAGAGATGACATCTCTTAAATGAAGCTTACAGATTTGTGTAAGAAAGAAA 831
 Oy 272 TyrluMetArgThrHisAsnArgValLysLeuAspGluAlaPheGluPheValLysGlnArg 291
 Db 832 AGACTTACTATATCTCAAACTTCAATTTCTGGGCGCAACTCCGAGACTATGAAAGAG 891
 Oy 292 ArgSerIleIleSerProAsnPheSerPheMetGlyGlnLeuLeuGlnPheGlnSerGln 311
 Db 892 ATT 894
 Oy 312 Val 312

Search completed: June 21, 2004, 12:35:00
Job time : 17.5253 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:25 ; Search time 46.8459 Seconds
(without alignments)
12204.244 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 1627
Sequence: 1 atggcccatgagatgatgg.....agaagataagaccagact 906

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=xip
-Q/cgnt 1/USPTO.spool.p/US10029345/runat.21062004.122816.4193/bsp.query.fasta.1.3278
-DB=SPRMBL.25 -OPT=fasta -SUFFIX=rspe -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345 @CGN 1.1 168 @runat.21062004.122816.4193 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPRMBL_25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phase: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp Vertebrate: *
14: sp Unclassified: *
15: sp_virus: *
16: sp_bacteriophage: *
17: sp_archaeal: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	1552	95.4	662	4 Q81VT8	Q81VT8 homo sapien

Result No.	Score	Query Match	Length	DB ID	Description
2	1545	95.0	665	4 Q96N49	Q96N49 homo sapien
3	1517	93.2	660	11 Q92OR2	Q92OR2 mus musculus
4	1514	93.1	677	11 Q99MG6	Q99MG6 mus musculus
5	1377.5	84.7	355	11 Q9AG16	Q9AG16 mus musculus
6	1312	80.6	616	4 Q8N5T1	Q8N5T1 homo sapien
7	1172.5	72.1	622	11 Q99MG5	Q99MG5 mus musculus
8	1036	63.7	300	11 Q9AG15	Q9AG15 mus musculus
9	989	60.8	625	4 Q86S88	Q86S88 homo sapien
10	985	60.5	665	11 Q7RS29	Q7RS29 mus musculus
11	891	54.8	206	11 Q85ZB4	Q85ZB4 mus musculus
12	653	40.1	143	4 Q96Q52	Q96Q52 homo sapien
13	563.5	39.6	657	5 Q8ST19	Q8ST19 caenorhabdi
14	472.5	29.0	367	13 Q7SZF3	Q7SZF3 brachydanto
15	469	28.8	483	11 Q8R3L3	Q8R3L3 mus musculus
16	465	28.6	382	13 Q7R2L9	Q7R2L9 gallus gall
17	450	27.7	606	5 Q8T1B8	Q8T1B8 caenorhabdi
18	447	27.5	383	13 Q80W48	Q80W48 fugu rubrip
19	440.5	27.1	368	4 Q8NFJ0	Q8NFJ0 homo sapien
20	436	26.8	382	13 Q7T2L8	Q7T2L8 brachydanto
21	430	26.4	362	13 Q803B2	Q803B2 brachydanto
22	427	26.2	411	4 Q13549	Q13549 homo sapien
23	423.5	26.0	378	13 Q91663	Q91663 xenopus lae
24	423.5	26.0	398	11 Q8BFV3	Q8BFV3 mus musculus
25	421.5	25.9	369	13 Q91790	Q91790 xenopus lae
26	419.5	25.8	369	13 Q90W58	Q90W58 xenopus lae
27	416	25.6	367	11 Q64193	Q64193 rattus bp.
28	413	25.4	318	11 Q80ZM1	Q80ZM1 mus musculus
29	413	25.4	367	11 Q63683	Q63683 rattus norv
30	411	25.3	439	5 Q81G35	Q81G35 caenorhabdi
31	403.5	24.8	353	13 Q42253	Q42253 gallus gall
32	403.5	24.8	360	13 Q7ZVL8	Q7ZVL8 brachydanto
33	403	24.8	436	11 Q99KC2	Q99KC2 mus musculus
34	397	24.4	452	11 Q8K1S9	Q8K1S9 mus musculus
35	397	24.4	452	11 Q7TNL7	Q7TNL7 mus musculus
36	382.5	23.5	354	4 Q8N4A4	Q8N4A4 homo sapien
37	347.5	21.4	411	5 Q9VSV1	Q9VSV1 drosophila
38	347.5	21.4	411	5 Q9VSV5	Q9VSV5 drosophila
39	325	20.0	303	4 Q9NSM1	Q9NSM1 homo sapien
40	306	18.8	177	11 Q9CSL5	Q9CSL5 mus musculus
41	287	17.6	369	5 Q44128	Q44128 caenorhabdi
42	285.5	17.5	476	5 Q46122	Q46122 drosophila
43	280.5	17.2	46	5 Q9VHV8	Q9VHV8 drosophila
44	272	16.7	290	5 Q86P14	Q86P14 drosophila
45	269	16.5	241	5 Q9VWV4	Q9VWV4 drosophila

ALIGNMENTS

RESULT 1
Q81VT8 PRELIMINARY; PRT; 662 AA.
ID Q81VT8
AC Q81VT8
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to dual specificity phosphatase 16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ductum;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC042101; AA042101.1; -
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPc; 1.


```

QY 121 CACATTTGGAGCCATTATATCAATGCTCCAGGCTTAAGAGGAGTTGCAACAG 180
DB 41 HieileuenglunalleaenilleaenCyserylLeuuectlysaargleuGln 60
QY 181 GACAAAGTGAATTAATCAGAGCTCATCCAGCATTCAGCGAAAGTAAGTTGACATTGAT 240
DB 61 AspLyValLeuilethrgluLeuileglnHieSerlalyshlysaValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATTGATTAAGATCAAAAGCTCCCAAGATGGCTCTCTCTTCA 300
DB 81 CysSerGlnlysaValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTTCTCACTGACTTCTGGGTAAGTGAAGAGGCTTCACTGTTCCCTG 360
DB 101 AspCyPheLeuThrValLeuLeuGlyLysleuGlnlySerPheAsnSerValHleu 120
QY 361 CTTCAGAGTGGGTTTGCTGAGTTCTCTCGTTGTTCCCTGGCCCTCTGTAAGGAAATCC 420
DB 121 LeuAlaGlylyPheAlaGlnPheSerArgCyPheProGlyLeuCySgluGlyLysSer 140
QY 421 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGCTTACCTGTGCAACATTTGGCCAAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCGCAGCGAGATGCTCTCAACAGAGCTGATA 540
DB 161 LeuIleLeuProAsnLeuTyrLeuGlyCySgluArgAspValLeuAsnlySgluLeuMet 180
QY 541 CACGAGATGGAGTGGTATGTGTTAAATGCCAGCATATCTGTCGAAGCTGACTTT 600
DB 181 GlnGlnAsnIlyIleGlyTyrValLeuAsnAlaSerAsnThrCySProlySProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTGTGAAATTTTG 660
DB 201 IleProGlnSerHiePheLeuArgValProValAsnSerPheCySgluLySleLeu 220
QY 661 CCGTGTGGACAAATGATGATGATTTGATGAGAAACAAAGCTCCATGATGATGTT 720
DB 221 ProTyrLeuAspLySerValAspPheIleGlnlysaIleAsnSerAsnGlyCyVal 240
QY 721 CTAGTGACTGTTAGTGGGATCTCCCGCTCCGCCACATCGCTATCGCTACATCATG 780
DB 241 LeuValHieCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGATGAGCATGTCTTATGATGAAGCTTACGATTTGTGAAGAAAAAGACTTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLySgluLySArgProThr 280
QY 841 ATATCTTCAAACTTCAATTTTCTGGGCAACTCTGAGACTTATGAGAAAGATTAGAAG 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnlySlySleLyAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302

```

```

RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
RT "MKP-7, a Novel Mitogen-activated Protein Kinase Phosphatase,
RT Functions as a Shuttle Protein.";
RL J. Biol. Chem. 276:39002-39011 (2001).
DR EMBL; AB052157; BAB47240.1; -.
DR HSSP; Q16828; MKP.
DR MGI; MGI:1917936; Dusp16.
DR GO; GO:0005737; C_cytoplasm; IDA.
DR GO; GO:0005634; C_nucleus; IDA.
DR GO; GO:0005515; F_protein binding; IPI.
DR GO; GO:0000188; P_inactivation of MAPK; IDA.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00782; DSPc; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ
SEQUENCE 660 AA; 72695 MW; DB609FCDADAAJ09 CRC64;

Alignment Scores:
Pred. No.: 3,6e-147 Length: 660
Score: 1517.00 Matches: 290
Percent Similarity: 98.34% Conservative: 7
Best Local Similarity: 96.03% Mismatches: 5
Query Match: 93.24% Indels: 0
DB: 11 Gaps: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x Q920R2 (1-660)
QY 1 ATGCCCATGACATGATTTGGAATCTCAATTTGTTACTGAGAGGTTGGTGCTCTCGGAA 60
DB 1 MetAlaHieGlnMetIleGlyThrGlnIleValThrGlnSerLeuValAlaLeuLeuGln 20
QY 61 AGTGAACGGAAAGAGCTGCTAATTTGATAGCGGCACTTTGTGNAATCAATATCATCC 120
DB 21 SerGlyThrGlnlysaIleLeuLeuIleAspSerArgProPheValIleTyrAsnThrSer 40
QY 121 CACATTTGGAGCCATTATATCAATCTGCTCAAGCTTATGAGCGAAGGTTGCAACAG 180
DB 41 HieileuenglunalleaenilleaenCyserylLeuuectlysaargleuGln 60
QY 181 GACAAAGTGAATTAATCAGAGCTCATCCAGCATTCAGCGAAAGTAAGTTGACATTGAT 240
DB 61 AspLyValLeuilethrgluLeuileglnHieSerlalyshlysaValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATTGATTAAGATCAAAAGCTCCCAAGATGGCTCTCTCTTCA 300
DB 81 CysAsnGlnlysaValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTTCTCACTGACTTCTGGGTAAGTGAAGAGGCTTCACTGTTCCCTG 360
DB 101 AspCyPheLeuThrValLeuLeuGlyLysleuGlnlySerPheAsnSerValHleu 120
QY 361 CTTCAGAGTGGGTTTGCTGAGTTCTCTCGTTGTTCCCTGGCCCTCTGTAAGGAAATCC 420
DB 121 LeuAlaGlylyPheAlaGlnPheSerArgCyPheProGlyLeuCySgluGlyLysSer 140
QY 421 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGCTTACCTGTGCAACATTTGGCCAAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCGCAGCGAGATGCTCTCAACAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuTyrLeuGlyCySgluArgAspValLeuAsnlySgluLeuMet 180

```

QY 541 CAGGAGATGAGATTGGTTATGCTTAAATGCCACTATACCTGTCCAAAGCCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyIleValIleuAsnAlaSerAnthrCysProIysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGGTGCTGCTGAGTGAATGACACTTTTGTGAGAAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIysIleLeu 220
QY 661 CCGTGGTTGACAATTCAGTATGATTCTTATGAGAAAGCAAGCCTTCAATGATGTT 720
DB 221 ProTyrLeuAspIysSerValAspPheIleGlyIleValIleuAlaSerAnthrCysVal 240
QY 721 CTAGGCACTGTTTACCTGGGATCTCCGCTCCGCACTGATGAGTATGCTTCCCTTACATG 780
DB 241 LeuIleHisCysLeuAlaGlyIleSerArgSerIleThrIleAlaIleAlaIleAlaIle 260
QY 781 AAGAGATGACATGCTTATGATGAGACTTACAGATTGTGAGAAAAGAAAGAAAGCTTACT 840
DB 261 LysArgMetAspMetSerIleAspGluAlaIleValArgPheValIleGlyIleValArgProThr 280
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGCTGATGAGAGAAAGAAAGCTTACT 900
DB 281 IleSerProAsnProAsnProMetGlyGlnLeuMetAspTyrGluIysThrIleAsnAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302

RESULT 4
Q99MG6 PRELIMINARY; PRT; 677 AA.
ID Q99MG6;
AC Q99MG6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2003 (TrEMBLrel. 17, Last sequence update)
DE Map kinase phosphatase-W A1 isoform.
GN DUSP16 OR 3830417M17R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=BALB/c;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RT Yoshikai Y.;
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
Activation in Macrophages.";
RL Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL: AF345951; AAK35052.1; -.
DR HSSP: Q16828; IMKP.
DR MGD: MGI:1917936; Dusp16.
DR GO: GO:0005737; C:cytoplasm; IDA.
DR GO: GO:0005634; C:nucleus; IDA.
DR GO: GO:0005515; F:protein binding; IPT.
DR GO: GO:0001088; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR001763; MAPK_phosph.
DR InterPro: IPR000843; MAPK_phosph.
DR InterPro: IPR000387; Tyr_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR PRINTS: PR01764; MAPKPHNTASE.
DR SMART: SM00450; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SEQUENCE 677 AA; 74550 MW; 886D5B7096CEC2FC CRC64;

Alignment Scores:
Pred. No.: 7.37e-147 Length: 677
Score: 1514.00 Matches: 289
Percent Similarity: 98.34% Conservative: 8
Best Local Similarity: 95.70% Mismatches: 5
Query Match: 93.05% Indels: 0
DB: 11 Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x Q99MG6 (1-677)

QY 1 ATGGCCCAATGAGATGATTTGAACTCAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGAA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValIleGlnSerLeuValAlaLeuGlu 20
QY 61 AGTGAACGAGAAAAGTCTGCTAATTTGATGAGCCGCACTTGTGGAATACATCATCC 120
DB 21 SerGlyThrGluValIleLeuIleAspSerArgProPheValGluTyrAnthrSer 40
QY 121 CACATTTGGAAGCCATTAAATCAACTGCTCCAACTTATGAGGAAAGGTTGCAACG 180
DB 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIysLeuMetLysArgIleGln 60
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGACTTACGCAAAACATTAAGTTGACATTGAT 240
DB 61 AspArgValLeuIleThrGluLeuIleGlnHisSerAlaLysHisValAspIleAsp 80
QY 241 TGCAGTCAGAAAGTTGTACTTTACATCAAAAGCTCCAAAGATGTGCTCTCTCTCA 300
DB 81 CysAsnGlnLysValAlaValIleValIleValIleValIleValIleValIleValIleVal 100
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAGAAAGCTTCACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuGluGlyIleLeuGlnIleValIleValIleValIleValIleVal 120
QY 361 CTTCAGGTGGTGTGCTGAGTTCTTCGCTGTTTCCCTGGCTCTGTGAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyIleValIleValIleVal 140
QY 421 ACTTACGCTCCATCTGATTTCTCACCTTGTGCTTACCTGTTGCAACATTGGCCAAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTCCCAATCTTATCTTCTGCTGCTGCAAGAGATGCTCTCAACAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuIleValIleValIleValIleValIleValIleValIleValIleVal 180
QY 541 CAGGAGATGAGATTGGTTATGCTTAAATGCCACTATACCTGTCCAAAGCCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyIleValIleuAsnAlaSerAnthrCysProIysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGGTGCTGCTGAGTGAATGACACTTTTGTGAGAAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIysIleLeu 220
QY 661 CCGTGGTTGACAATTCAGTATGATTCTTATGAGAAAGCAAGCCTTCAATGATGTT 720
DB 221 ProTyrLeuAspIysSerValAspPheIleGlyIleValIleuAlaSerAnthrCysVal 240
QY 721 CTAGGCACTGTTTACCTGGGATCTCCGCTCCGCACTGATGAGTATGCTTCCCTTACATG 780
DB 241 LeuIleHisCysLeuAlaGlyIleSerArgSerIleThrIleAlaIleAlaIleAlaIle 260
QY 781 AAGAGATGACATGCTTATGATGAGACTTACAGATTGTGAGAAAAGAAAGAAAGCTTACT 840
DB 261 LysArgMetAspMetSerIleAspGluAlaIleValArgPheValIleGlyIleValArgProThr 280
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGCTGATGAGAGAAAGAAAGCTTACT 900
DB 281 IleSerProAsnProAsnProMetGlyGlnLeuMetAspTyrGluIysThrIleAsnAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302

DR PROSITE; PS50056; TYR PHOSPHATASE 2; 1.
 DR PROSITE; PS50054; TYR PHOSPHATASE_DUAL; 1.
 KW Hypothetical protein.
 FT NON TER
 SQ SEQUENCE 616 AA; 67636 MW; 2CB0B14482F2AD72 CRC64;

Alignment Scores:

Pred. No.: 4, 676-126
 Score: 1312.00
 Percent Similarity: 99.60%
 Best Local Similarity: 99.21%
 Query Match: 80.64%
 DB: 4
 Gaps: 0

US-10-029-345a-108_COPY_538_1443 (1-906) x Q99MG5 (1-616)

```

QY 148 TCCTCAAGCTTATTAAGCGAAGTTGCGACAGACGAAAGTTGTTAATTACAGAGCTCATC 207
DB 1 CysSerLysLeuMetLysArgArgLeuGlnGlnAspLysValLeuIleThrGluLeuIle 20
QY 208 CAGCATTCAGCGAAGAACTAAGTTGACATTGATTCAGATTCAGAGAGTTGTTAGCAT 267
DB 21 GlnHisSerAlaLysHisLysValAspIleAspCysSerGlnLysValValValLysAsp 40
QY 268 CAAAGCTCCCAAGATGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 327
DB 41 GlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrValLeuGlu 60
QY 328 AAACGTGAGAGAGAGCTTCAACTCTGTTCACTGCTTGCAGAGTGAGTTGCTGAGTTCTCT 387
DB 61 LysLeuGlnLysSerPheAsnSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSer 80
QY 388 CGTTGTTCCCTGCGCTCTGTGAAGAAATTCACCTGATCCCTTACCTGATTCCTGAC 447
DB 81 ArgCysPheProGlyLeuGlyCysGlyLysSerThrLeuValProThrCysLysSerGln 100
QY 448 CCTTGCTTACCTGTTGCGCAAGATGGGCGAACCCGAAATTTCTCCATCTTATCTGGC 507
DB 101 ProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuIleLeu 120
QY 508 TGCCAGCAGAGTTCCTCAACAGAGAGCTGATCAGCAGATGGAGATGGATTTGTTATGTTA 567
DB 121 CysGlnHisArgAspValLeuAsnLysGluLeuMetGlnGlnAsnGlyIleGlyThrValLeu 140
QY 568 AATGCCAGCTATACCTGTCGAAGCCTGACTTATATCCCGAGCTCATTTCTGCGGTG 627
DB 141 AsnAlaSerAsnThrCysProLysProAspPheIleProGluSerHisPheLeuArgVal 160
QY 628 CCTGTGAATGACAGCTTTTGTGAAGAAATTTGCCGTGTTGAGCAATGAGTATGATTTTC 687
DB 161 ProValAsnAspSerPheCysGluLysIleLeuProThrLeuAspLysSerValAspPhe 180
QY 688 ATTGAAGAAAGCAAAAGCCTCAATGATGTGTTCTAGTGCAGCTGTTAGCTGGATCTCC 747
DB 181 IleGluLysAlaLysAlaSerAsnLysCysValLeuValHisCysLeuAlaGlyIleSer 200
QY 748 CGCTCCGACCATGCGATGCGCTCAATCATGTAAGAGATGAGCAATGCTTTAGATGA 807
DB 201 ArgSerAlaThrIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 220
QY 808 GCTTCAAGATTTGTGAAGAAAGAAAGAAAGCTACTATATCTCCAAATCTTCAATTTTCTGG 867
DB 221 AlaTyrArgPheValLysGluLysArgProThrIleSerProAsnPheAsnPheLeuGly 240
QY 868 CAACCTCTGAGCTTGAAGAGAGATTAAGAACCAAGCT 906
DB 241 GlnLeuLeuAspLysArgLysLysIleLysAsnGlnThr 253

```

RESULT 7

Q99MG5 PRELIMINARY; PRT; 622 AA.
 AC Q99MG5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Map kinase phosphatase-M A2 isoform.
GN DUSP16 OR 3830417M17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.;
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages."
RL Mol. Cell. Biol. 20:6999-7009 (2001).
DR EMBL; AF345952; AAK35053.1; -.
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IDA.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro; IPR000340; P:inactivation of MAPK; IDA.
DR InterPro; IPR008343; MAPK phosphatase.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANASE 3; 1.
DR PROSITE; PS00383; TYR PHOSPHATASE 1; 1.
DR PROSITE; PS50056; TYR PHOSPHATASE 2; 1.
DR PROSITE; PS50054; TYR PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ SEQUENCE 622 AA; 68672 MW; 6C0CA4BE909B98 CRC64;

```

Alignment Scores:

Pred. No.: 1, 1e-111
 Score: 1172.50
 Percent Similarity: 80.13%
 Best Local Similarity: 77.48%
 Query Match: 72.07%
 DB: 11
 Gaps: 1

US-10-029-345a-108_COPY_538_1443 (1-906) x Q99MG5 (1-622)

```

QY 1 ATGGCCCATGAGATGATTTGAAGCTCAATTTGTTACGAGGTTGGCTGTGCTGAGAA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrIleSerLeuValAlaLeuGln 20
QY 61 AGTGAAGCGAAGAAAGTCTCTTAATGATAGCCGCGCATTTGTGGAATATCAATCATCC 120
DB 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluLysAsnThrSer 40
QY 121 CACATTTTGAAGCCATTAAATATCAATGCTGCCAAGCTTATGAAGCAAGGTTGACAG 180
DB 41 HisIleLeuGlnAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 60
QY 181 GACAAAGTGTAAATTAAGAGCTCATTCGCAAGCTTCAAGCAACATTAAGTTGACATTGAT 240
DB 61 AspArgValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCAAGTCAAGAGTTGATTTAGATTCAGATCAAGCTCCGAAGATGTTGCTCTCTTCA 300
DB 81 CysAsnGlnLysValValValLysArgPheIleSerGlnAspValGlySerLeuSerSer 100
QY 301 GACTGTTTCTCACTGATCTTGTGGTAACTGGAAGAGAGCTTCACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnArgSerPheAsnSerValHisLeu 120

```

```

QY 361 CTTCAGAGTGGGTTTGCTGAGATCTCTGTTGTTTCCCTGGCCTCTGTGAAGAAATCC 420
    |||||
Db 121 LeuAla----- 122
QY 421 ACTTACTCCCTACTGCTGATTTCTACGCTTGCTTACTGTGCAACATTGGCCCAACC 480
    |||||
Db 122 ----- 122
QY 481 CGAATTTCTCCCAATCTTTATCTTGCTGCTGCAGGAGATGTCTCAACAAGAGCTGATA 540
    |||||
Db 123 -----AspLeuMet 125
QY 541 CAGCAGATGGAGTTGGTTATGTGTTAAATGCCAGTATACCTGTCCAAAGCCTGACTTT 600
    |||||
Db 126 GINGlhaAnnglylleGlyTyValLeuAAnhAAsenAnThCySProlYsProlAsPhe 145
    |||||
QY 601 ATCCCCGAGTCTCATTTCTGCGGTGCTGTGATGATGACAGCTTTTGAGAAATTTTG 660
    |||||
Db 146 lLeProluSerhAAsPheLeuArgValProValAAsnAsPheSerPheCySGluYsIleLeu 165
    |||||
QY 661 CCGTGGTGGACAAATCATAGTATGATTTTATGAGAAACAAAGCCTCAATGATGTT 720
    |||||
Db 166 ProTrpLeuAspYsSerValAsPheIleGluYsAlAsPheAAsenAnGlyCyVal 185
    |||||
QY 721 CTAGTGCATGTTAGCTGGGATCTCCGCTCCGCGCACATCGCTATGCGCTACATGATG 780
    |||||
Db 186 LeuIleAAsCySeuAAlaglylIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 205
    |||||
QY 781 AAGAGATGACATGCTTTAGATGAGATGAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
    |||||
Db 206 LyAAsGheAAsPheMetSerleuAAsPheAlaTyrlAAsGheValYsGluYsArgProThr 225
    |||||
QY 841 ATATCTTCAAACTTCAATTTTCTGGGCAACTCTGTGACATATGAGAAAGATTAAGAAC 900
    |||||
Db 226 lIleSerProAsnAAsPheAAsPheMetGlyGluLeuMetAspTyrlGluYsThrIleAAsn 245
    |||||
QY 901 CAGACT 906
    |||||
Db 246 GluThr 247

```

```

DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSRP; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00206; RHODANSE_3; 1.
DR PROSITE; PS00363; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW HydroLase; Kinase.
SQ SEQUENCE 300 AA; 33624 MW; 4C61846ACDF0F456 CRC64;

Alignment Scores:
Pred. No.: 1,09e-97 Length: 300
Score: 1036.00 Matches: 217
Percent Similarity: 72.20% Conservative: 9
Best Local Similarity: 69.33% Mismatches: 20
Query Match: 63.68% Indels: 67
DB: 11 Gaps: 2

US-10-029-345a-108_COPY_538_1443 (1-906) x Q9AG15 (1-300)

QY 1 ATGCCCATGAGATGATGGAATCTCAATTTGTTACTGAGAGTTGGCTGTCTGGAA 60
    |||||
Db 1 MetAlAAsGluMetIleGlyThrGlnIleValThrGluSerLeuValAlaLeuGlu 20
    |||||
QY 61 AGTGAAGGAAAGAGTGGCTTAATGATAGCCGGCCATTTGTGAATACATACATCC 120
    |||||
Db 21 SerGlyThrGluYsValLeuLeuIleAsPheArgProPheValGluTyraenThrSer 40
    |||||
QY 121 CACATTTGGAAGCCATTAATATCACTGCTCCAAAGCTTATGAGAAAGTTGCAACAG 180
    |||||
Db 41 HieIleGluAlaIleAAsenIleAAsenIleAAsenIleAAsenIleAAsenIleAAsen 60
    |||||
QY 181 GACAAAGTTAATTAACAAGCTCATCAGACATTCAGCGAAACATTAAGTTGACATTGAT 240
    |||||
Db 61 AsPArgValLeuIleThrGluLeuIleGlnHIsSerAlaIleHIsIleYsValAsPheIleAsp 80
    |||||
QY 241 TGCAGTCAGAGGTTGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 300
    |||||
Db 81 CyAAsGlnYsValAlaValAlaValAlaValAlaValAlaValAlaValAlaValAlaVal 100
    |||||
QY 301 GACTGTTTCTCACTACTTCTGAGTAACTGAGAAAGCTTCAACTCTGTTCACCTG 360
    |||||
Db 101 AsPArgPheLeuThrAlaLeuLeuGlyYsIleGluYsArgSerPheAAsenIleAAsen 120
    |||||
QY 361 CTTCAGAGTGGGTTGCTGAGTCTCTGCTGTTTCCCTGGCCTCTGTGAGAGAAATCC 420
    |||||
Db 121 LeuAla----- 122
QY 421 ACTTACTCCCTACTGCTGATTTCTACGCTTGCTTACCTGTTCACAAATTTGGCCCAACC 480
    |||||
Db 122 ----- 122
QY 481 CGAATTTCTCCCAATCTTTATCTTGCTGCTGCAGGAGATGTCTCAACAAGAGCTGATA 540
    |||||
Db 123 -----AspLeuMet 125
QY 541 CAGCAGATGGAGTTGGTTATGTGTTAAATGCCAGTATACCTGTCCAAAGCCTGACTTT 600
    |||||
Db 126 GINGlhaAnnglylleGlyTyValLeuAAnhAAsenAnThCySProlYsProlAsPhe 145
    |||||
QY 601 ATCCCCGAGTCTCATTTCTGCGGTGCTGTGATGATGACAGCTTTTGAGAAATTTTG 660
    |||||
Db 146 lLeProluSerhAAsPheLeuArgValProValAAsnAsPheSerPheCySGluYsIleLeu 165
    |||||
QY 661 CCGTGGTGGACAAATCATAGTATGATTTTATGAGAAACAAAGCCTCAATGATGTT 720
    |||||
Db 166 ProTrpLeuAspYsSerValAsPheIleGluYsAlAsPheAAsenAnGlyCyVal 185
    |||||
QY 721 CTAGTGCATGTTAGCTGGGATCTCCGCTCCGCGCACATCGCTATGCGCTACATGATG 780
    |||||
Db 186 LeuIleAAsCySeuAAlaglylIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 205
    |||||

```

QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTGGAAGAA----- 828
 Db |||||
 QY 206 LyarGwetaSpMetSerLeuSpGluAlaYrAArgGlnYrProThrAspArgAla 225
 QY 829 -----AAAAGCCTACTATATCTCCAACTTCATTTTGGGC 867
 Db 226 AspSerAArgThrGlyMetLeYrSarGAlaProLeuYrSerSerLeuAsn-AlaGluAla 245
 QY 868 CAATCTCTGACTGATGAGAAAGATTAAGAACGAGA 904
 Db 245 aaAlaYrSTrPLeuLeuGluAArgAlaLeuCysArgArg 257

RESULT 9

Q86S88 PRELIMINARY; PRT; 625 AA.
 ID 086S88
 AC 086S88
 DT 01-JUN-2003 (TRENBLREL. 24, Created)
 DT 01-JUN-2003 (TRENBLREL. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Similar to dual specificity phosphatase 8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NC NCI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAUSBERG R., and Astrocytoma;
 RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC045110; AA045110.1; -
 DR GO; GO:0017017; F.MAP kinase phosphatase activity; IEA.
 DR GO; GO:0006470; P.protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR002965; P-rich exten.
 DR InterPro; IPR001763; Rhodanese-like.
 DR Pfam; PF00782; DSPC; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR PRINTS; PR01217; PRICHEXTENSN.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANES_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 SQ SEQUENCE 625 AA; 65826 MW; C7C808407B724FFC CRC64;

Alignment Scores:

Pred. No.: 8.84e-93 Length: 625
 Score: 989.00 Matches: 190
 Percent Similarity: 81.40% Conservative: 55
 Best Local Similarity: 63.12% Mismatches: 54
 Query Match: 60.79% Indels: 2
 DB: 4 Gaps: 2

US-10-029-345a-108_COPY_538_1443 (1-906) x Q86S88 (1-625)

QY 1 ATGGCCCATGAGATATGGAATCTAATTTGTT---ACTGAGGTTGGTGGCTGCTGCT 57
 Db |||||
 QY 1 MetAlaGlyAspArgLeuProArgLysValMetCasPalalysLysLeuAlaSerLeu 20
 QY 58 GAAAGTGGACGGAAGAAAGTGTCTAATGATGACCGGCACTTTGTGGAATACATACA 117
 Db |||||
 QY 21 ArgGlyGlyProGlyGlyProLeuValLleAspSerAArgSerPheValGluThrAsnSer 40
 Db 118 TCCGACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGACGAAAGGTTGCA 177
 QY 41 TrpHisValLeuSerSerValaLleuLleCysCysSerLysLeuValLysArgArgLeuGln 60
 QY 178 CAGGACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATT 237
 Db 61 GlnGlyLysValThrLleAlaGluLeuLleGlnProAlaAlaArgSerGlnValGluAla 80

QY 238 GATTGACATGAGAGGTTGAGTTAGATTACATCAAGATCCCAAGATGTGCTCTCTCT 297
 Db |||||
 QY 81 ThrGluProAlaAspValValValYrAspGlnSerThrAArgAlaSerValLeuAla 100
 QY 298 TCAGACTGTTTCTCAGTCTACTGTAACCTGGAAGAAAGACTTCATCTGTTTAC 357
 Db |||||
 QY 101 AlaAspSerPheLeuSerLleLeuLeuSerLysLeuAspGlyCysPheAspSerValAla 120
 QY 358 CTGCTTGACGGGAGGTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGGAAGAAA 417
 Db |||||
 QY 121 IleLeuThrGlyGlyPheAlaThrPheSerSerCysPheProLysLeuGlySerGluGlyLys 140
 QY 418 ---TCACCTTAGTCCCTACCTGATCTTCACCTGCTTACCTTCTTCCAACTGGG 474
 Db |||||
 QY 141 ProAlaAlaLeuLeuProMetSerLeuSerGlnProCysLeuProValProSerValGly 160
 QY 475 CCAACCCGAATTTCTCCCAATCTTTATCTTGCTGCCAGCAGATGCTCTCAACAGAG 534
 Db |||||
 QY 161 LeuThrArgLleLeuProHisLeuYrLeuGlySerGlnLysAspValLeuAsnLysAsp 180
 QY 535 CTGATACAGACAGATGGAGATTGTTATGTTAAATGACAGTAACTGCTTCCAAAGCTT 594
 Db |||||
 QY 181 LeuMetThrGlnAsnGlyLleSerYrValLeuAsnAlaSerAsnSerCysProLysPro 200
 QY 595 GACTTATATCCCGAGTCTCATTTCTGCGTGTGCTGTAATGACAGCTTGTGGAAGAA 654
 Db |||||
 QY 201 AspPheLleCysGlnSerAArgPheMetAArgValProLleAsnAspAenYrCysGluLys 220
 QY 655 ATTTTCCCTGCTGTTGACAAATCAATGATTTCAATTGAGAAAGCAAGCTTCATGCA 714
 Db |||||
 QY 221 LeuLeuProTrpLeuAspLysSerLleGluPheLleAspLysAlaLysLeuSerCys 240
 QY 715 TGTGTTCTAGTCACTGTTAGCTGGATCTCCCGCTGCCCACTGCTCTATGCTTAC 774
 Db |||||
 QY 241 GlnValLleValHisCysLeuAlaGlyLleSerAArgSerValThrLleAlleAlaYr 260
 QY 775 ATCATGAGAGAGATGACATGCTTTAGATGAGACTTATGAGATTGTGTAAGAAAAAGA 834
 Db |||||
 QY 261 IleMetLysThrMetGlyMetSerSerAspAspAlaLtyAArgPheValLysAspArgArg 280
 QY 835 CCTACTATATCTCCAACTTCATTTTCTGGGCAACTCTGACTATGAGAAAGATT 894
 Db |||||
 QY 281 ProSerLleSerProAsnPheAsnPheLeuGlyGlnLeuGluYrGluAArgSerLeu 300
 QY 895 AAG 897
 Db |||||
 Db 301 Lys 301

RESULT 10

Q77S29 PRELIMINARY; PRT; 665 AA.
 ID 077S29
 AC 077S29
 DT 01-OCT-2003 (TRENBLREL. 25, Created)
 DT 01-OCT-2003 (TRENBLREL. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Dusp8 protein
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAUSBERG R., and Astrocytoma;
 RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC045110; AA045110.1; -
 DR GO; GO:0006470; P.protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR002965; P-rich exten.
 DR InterPro; IPR001763; Rhodanese-like.
 DR Pfam; PF00782; DSPC; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR PRINTS; PR01217; PRICHEXTENSN.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANES_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 SQ SEQUENCE 665 AA; 68826 MW; C7C808407B724FFC CRC64;

QY 121 CACATTTGGAGCCATTAATATCACTGCTCCAAAGCTTATGAGCGAAGTTGCAACAG 180
 Db |||||
 QY 41 HistileugluhalaileansnleamscysserlyleuMethylsArgargleugln 60
 Db |||||
 QY 181 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAGCAATAGATTGACATTGAT 240
 Db |||||
 QY 61 Asplyvalleullethrgileuileglnhlserralyshlsylsvalaspileasp 80
 Db |||||
 QY 241 TGCAGTCGAGAGGTTGATGATTAAGATCAAGCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
 Db |||||
 QY 81 CysasnglnlyvalValValValValAspInserSerInAspValAlasylserleuSer 100
 Db |||||
 QY 301 GACTGTTTCTACGCTGCTCTGCGTAACTGAGAGAGAGCTTCAACTGTCACCTGCACTG 360
 Db |||||
 QY 101 AspCyspheleuThrValleuLeuGlyLysleuGlnLysSerPheAsnSerValHisleu 120
 Db |||||
 QY 361 CTTCAGAGTGGGTTGCTGAGTTCTCGTTGTTCCCTGGCTCTGTGAGAGAAATCC 420
 Db |||||
 QY 121 LeuAlaglyGlyPheAlaglnPheSerArgCysPheProGlyLeuCysgluGlyLysSer 140
 Db |||||
 QY 421 ACTCTAGTCCCTTACCTGCTCATTTCTCAGCTTGTCTTACCTGTTGCCAATGGGCAACC 480
 Db |||||
 QY 141 ThrleuValProthrcysIleSerGlnProCysleuProValAlaAsnIleGlyProthr 160
 Db |||||
 QY 481 CGAATTTTCCCAATCTTATCTTGCTGCTCCAGCGAGATGTCCTCAACAG 531
 Db |||||
 QY 161 ArgileuProAsnleuThryleuGlyCysgluAlaspileuValLeuValLys 177

RESULT 12
 Q96QS2 PRELIMINARY; PRT; 143 AA.

AC Q96QS2; 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Truncated MAPK phosphatase 7.
 GN MKP7.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Montpetit A., Bolly G., Simmet D.;
 RT "A detailed transcriptional map of the chromosome 12p12 tumor
 suppressor locus."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY038927; AAK69770.1; -
 DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PRO1764; MAPKPHPTASE.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS02026; RHODANASE 3; 1.
 SO SQUIDNCE 143 AA; 16077 MW; 5213A213AA75974 CRC64;

Alignment Scores:
 Pred. No.: 2.68e-58
 Score: 653.00
 Percent Similarity: 70.35%
 Best Local Similarity: 70.35%
 Query Match: 40.14%
 Length: 143
 Matches: 140
 Conserves: 0
 Mismatches: 3
 Indels: 56
 Gaps: 1

US-10-029-345a-108_copy_538_1443 (1-906) x Q96QS2 (1-143)

QY 1 ATGGCCCATGAGATGATTTGGAAGTCAATATTTACTGAGAGTTGGTCTGCTGGA 60
 Db |||||
 QY 1 MechahiselmethyllethrglnleuValThrGluValAlaLeuLeuGlu 20
 Db |||||
 QY 61 AGTGAACGGAAGAAAGTGTGTAATGATAGCGCGCATTTGTGGAATCAATACATCC 120
 Db |||||

Db |||||
 QY 21 SerGlyThrClnlyValleuLeulleaspeArpProPheValGlnlyrAsnThrSer 40
 Db |||||
 QY 121 CACATTTGGAGCCATTAATATCACTGCTCCAAAGCTTATGAGCGAAGTTGCAACAG 180
 Db |||||
 QY 41 HistileugluhalaileansnleamscysserlyleuMethylsArgargleugln 60
 Db |||||
 QY 181 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAGCAATAGATTGACATTGAT 240
 Db |||||
 QY 61 Asplyvalleullethrgileuileglnhlserralyshlsylsvalaspileasp 80
 Db |||||
 QY 241 TGCAGTCGAGAGGTTGATGATTAAGATCAAGCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
 Db |||||
 QY 81 CysasnglnlyvalValValValValAspInserSerGlnAspValAlasylserleuSer 100
 Db |||||
 QY 301 GACTGTTTCTACGCTGCTCTGCGTAACTGAGAGAGAGCTTCAACTGTCACCTGCACTG 360
 Db |||||
 QY 101 AspCyspheleuThrValleuLeuGlyLysleuGlnLysSerPheAsnSerValHisleu 120
 Db |||||
 QY 361 CTTCAGAGTGGGTTGCTGAGTTCTCTCGTTGTTCCCTGGCTCTGTGAGAGAAATCC 420
 Db |||||
 QY 121 LeuAlagly--Ala-----
 Db |||||
 QY 421 ACTCTAGTCCCTTACCTGCTCATTTCTCAGCTTGTCTTACCTGTTGCCAATGGGCAACC 480
 Db |||||
 QY 124 -----
 Db |||||
 QY 481 CGAATTTTCCCAATCTTATCTTGCTGCTCCAGCGAGATGTCCTCAACAGAGCTGATTA 540
 Db |||||
 QY 125 -----
 Db |||||
 QY 541 CAGCGAATGGAGTGTATGTTATGTTAAATGCCAGTATACCTGTCGAAGCC 593
 Db |||||
 QY 126 lalaglnlThrAspThrleuCysVallyscysgluGlnlyrleuSerlyAla 143

RESULT 13
 Q8ST19 PRELIMINARY; PRT; 657 AA.

AC Q8ST19; 01-JUN-2002 (TRENBLREL. 21, Created)
 DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Hypothetical protein F08B1.1a.
 GN F08B1.1.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peleoderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Bristol N2;
 RX MEDLINE=9069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Bristol N2;
 RC Chisese S.;
 RT "The sequence of C. elegans cosmid F08B1.";
 RL Submitted (Apr-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U23178; AAK68295.1; -
 DR HSSP; Q16828; IMKP.
 DR WormPep; F08B1.1a; CE27918.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.
 DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.

DR GO:0006470; P:Protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR000340; DS phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-1like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00195; DSPC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS50206; RHODANSE_3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR Hypothetical protein: Hydrolase.
 KW SEQUENCE 657 AA; 71002 MW; A0D9153DE6326B43 CRC64;

Alignment Scores:

Pred. No.:	6,11e-49	Length:	657
Score:	563.50	Matches:	131
Percent Similarity:	61.54%	Conservative:	61
Best Local Similarity:	41.99%	Mismatches:	87
Query Match:	34.63%	Indels:	33
		Gaps:	10

US-10-029-345A-108_COPY_538_1443 (1-906) x Q8ST19 (1-657)

QY 28 ATTGTTCTGAGAGGTTGGCTCTCTGGAAGTGAACGAAAGTCTGCTAATT 87
 DB 9 ILSerThrCysglLeuAlaLeuIleArgGluAlaProAspThrThrLeuVal 28
 QY 88 GATAGCGCGCATTTGGATTCATATACATATACATATTTGGAGCATTAATACAC 147
 DB 29 AspCysArgGlyPheThrGluTyrAsnGluSerHisValArgHisSerMetAlaAlaPhe 48
 QY 148 TGCTCCAGCTTAGAGCGAAGGTTGCCAAGAGCAAGTGTG-----TTAATT 195
 DB 49 PheSerIleLeuIleArgGluArgLeuPheGluAlaAsnIleLeuAspAlaPheCysLeuIle 68
 QY 196 ACAGAGCTCATCCAGCATTGAGC-----AAACATAAGTTGACATTGAT 240
 DB 69 HisGlnLeuMetSerCysSerSerGlyCysThrIleMetAspGluIleAspLeu--- 87
 QY 241 TGCAGTGAAGAGTTGAGTTTACGATCA-----AGCTCCCAAGAT 282
 DB 88 -----ValLeuTyrAlaGluGluAspLysProArgIleAsnLysArgArg 102
 QY 283 GTTGCTCTCTCTC-----TCTTCAGACTGTTTCTCACTGACTTGGGTA 330
 DB 103 IleAlaSerCysAsnAlaProGluSerThrAlaLysIleMetArgValLeuArgGluArg 122
 QY 331 CTGGAGAGAGC-----TTCAACTGTTCACCTGCTTCAGAGTGGTTGCTGAGTTTC 384
 DB 123 LeuGluAspThrAspLysPheArgSerValMetValLeuGluGlyIlePheLysGlnPhe 142
 QY 385 TCTCGTGTCTTCCTGCTGCTGCTGAGAGAAATCC-----ACTCAGTCCCTACTCTGC 438
 DB 143 AlAGlnGlnIleTyrProGlnLeuCysGluSerSerGluIleMetThrArgLeuProGlnSer 162
 QY 439 ATTGTTGAGCTTGTCTTA-----CCTGTTGCCAACAATTGGGCAACCCGAATTTCTTCCC 492
 DB 163 LeuSerGlnProCysLeuSerGlnProThrGlyAsp---GlyIleThrLeuIleThrPro 181
 QY 493 AATCTTATCTTGGCTGCGCAGCGAGATGCTCTCAACAAGAGCTGATACAGCAATGGG 552
 DB 182 AsnIleTyrLeuGlySerGlnIleAspSerLeuAspGluThrMetLeuAspAlaLeuAsp 201
 QY 553 ATGGTATGATGGTAAATGCGACTATACCTGTCGCAAGGCTGACTTATCCCGAG--- 609
 DB 202 IleSerValValIleAsnLeuSerMetThrCysProLysSerValCysIleLysGluAsp 221
 QY 610 TCTCATTTCTCGCTGCTGCTGCTGATGATGACAGCTTTTGTGAGAAATTTTCCGTGTTG 669
 DB 222 LysAsnPheMetArgIleProValAsnAspSerTyrGlnIleLysLeuSerProTyrPhe 241

QY 670 GACAATCAGTATTTCAATTGAGAGAGAAAGCCCAATGATGATGTTTCTAGTCAC 729
 DB 242 PrometAlaTyrGlnPheLeuGluLysCysArgArgAlaIleLysLysCysLeuIleHis 261
 QY 730 TGTTAGCTGAGATCTCCGCTCCGCCACCATGCTATGCTGCTATCATGAGAGATG 789
 DB 262 CysLeuAlaGlyIleSerThrArgSerProThrLeuAlaIleSerTyrIleMetArgTyrMet 281
 QY 790 GACATGCTTTAGATGAGAGCTTACAGATTGTGTAAGAAAGAAAGACCTACTATATCTCCA 849
 DB 282 LysMetGlySerAspAspAlaTyrArgTyrValLysGluArgArgProSerIleSerPro 301
 QY 850 AACTCAATTTTCTGGGCCCAACTCTGAGACTAGAG 885
 DB 302 AsnPheAsnPheMetGlyGlnLeuLeuGluTyrGln 313

RESULT 14

ID	Q7SZF3	PRELIMINARY;	PRT;	367 AA.
----	--------	--------------	------	---------

AC Q7SZF3; PRELIMINARY; PRT; 367 AA.
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Body;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins T.B., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haile F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Tohilyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywicki M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.U., Matra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Body;
 RA Strausberg R.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC052477; AAH52477.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 367 AA; 40538 MW; F89997A415DAD6AC CRC64;

Alignment Scores:

Pred. No.:	1.28e-39	Length:	367
Score:	472.50 <td>Matches:</td> <td>104</td>	Matches:	104
Percent Similarity:	56.86%	Conservative:	66
Best Local Similarity:	34.78%	Mismatches:	110
Query Match:	29.04%	Indels:	19
		Gaps:	6

US-10-029-345A-108_COPY_538_1443 (1-906) x Q7SZF3 (1-367)


```
Db      274 LeuSerSerPheIysGlnAsnHisGlyAsnLeuCysAspAsnSerLeuGlnLeuGlnIu 293
QY      418 -----TCCACTCTAGTCCTTACCTGCATTCT 444
Db      294 CysArgGluValGlyGlyAlaSerAlaAlaSerSerMetLeuProGln-----Ser 311
QY      445 CAGCCTTGCTTACCT---GTTGCCACATTGGGCCAACCCGAATTCTTCCCAATCTTTAT 501
Db      312 ValProThrThrProAspIleGluAsnAlaGluLeuThrProIleLeuProPheLeuPhe 331
QY      502 CTGGCTGCGCAGCGAGATGTCCTCCACAGAGAGCTGATACAGACAAATGGGATTGTTAT 561
Db      332 LeuGlyAsnGlnAspAlaGlnAspLeuAspThrMetGlnArgLeuAsnIleGlyTyr 351
QY      562 GTGTTAATGCCAGC-----TATACCTGTCCAAAGCCTGACTTTATCCC 606
Db      352 ValIleAsnValThrThrHisLeuProLeuTyrHisTyrGluYsglyLeuPhe----- 369
QY      607 GAGTCTCATTTCCGCTGCTGCTGTGATGACAGCTTTTGTGAGAAATTTGCGCTGG 666
Db      370 -----AsnTyrIysArgLeuProAlaThrAspSerAsnIysGlnAsnLeuArgGlnTyr 387
QY      667 TTGGACAATCAGTAGATTTCATTGAGAAACAAAGCCTCCATGGATGTTCTTAGTG 726
Db      388 PheGluGlnAlaPheGluPheIleGluGlnAlaHisGlnCysGlyYsglyLeuLeuIle 407
QY      727 CACTGTTTACGTGGGATCTCCGCTCCGCCACCATCGCTATCGCTTACATCATGAAGAG 786
Db      408 HisCysGlnAlaGlyValSerArgSerAlaThrIleValIleAlaTyrLeuMetLysHis 427
QY      787 ATGGACATGCTCTTAGATGAAGCTTACAGATTGTGAAGAAAGAAAGACCTACTATATCT 846
Db      428 ThrArgMetThrMetThrAspAlaTyrIlyAspHeValYsglyYsArgProIleIleSer 447
QY      847 CCAAACTTCAATTTCTGGGCCCAACTCTGACTATGAGAGAGATTAAAGAC 900
Db      448 ProAsnLeuAsnPheMetGlyGlnLeuLeuGlnPheGluGlnAspLeuAsnAsn 465
```

Search completed: June 21, 2004, 12:40:25
Job time : 53.8459 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_n2p model

Run on: June 21, 2004, 12:28:21 ; Search time 43.4105 Seconds
(without alignments)
11793.817 Million cell updates/sec

Title: US-10-029-345a-108_COPY_538_1443

Perfect score: 1627
Sequence: 1 atggccacatgagatgcttgg.....agaagattaagaccagact 906

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlp
-Q/cg2.1/USPTO.spool.p/US10029345/runat.21062004.122814.4167/app.query.fasta.1.3278
-DB=A.GeneSeq.29Jan04 -QPMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345.QCEN.1.1_128.qrunat.21062004.122814.4167 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_GeneSeq.29Jan04.*

1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1564	96.1	665	5	ABR52381
2	1564	96.1	665	5	ABR52407
3	1552	95.4	665	4	AAE04834
4	1552	95.4	665	4	AAU09016
5	1552	95.4	665	5	AAU09156
6	1552	95.4	665	5	AAU09946
7	1552	95.4	665	5	AAU75789
8	1552	95.4	665	5	ABR97946
9	1552	95.4	665	5	AAU79929
10	1552	95.4	665	5	ABR97291

11	1552	95.4	666	4	ABR20325	ABR20325 Human pro
12	1545	95.0	665	5	ADA54744	ADA54744 Human pro
13	1544	94.9	665	5	AAU79161	AAU79161 Human dua
14	1544	94.9	665	5	ABR52352	ABR52352 Protein r
15	1544	94.9	672	4	AAU25744	AAU25744 Human pro
16	1542	94.8	665	5	AAU79162	AAU79162 Human dua
17	1537	94.5	302	5	ABR52425	ABR52425 Protein r
18	1526.5	93.8	664	5	ABR52424	ABR52424 Protein r
19	1482	91.7	660	5	ABR52385	ABR52385 Protein r
20	988	60.7	625	5	ABR52382	ABR52382 Protein r
21	988	60.7	625	5	ABR52350	ABR52350 Protein r
22	988	60.7	625	5	ABR52340	ABR52340 Human dua
23	985	60.5	663	2	AAW29150	AAW29150 Dual -spec
24	985	60.5	663	5	ABR52351	ABR52351 Protein r
25	742.5	45.6	375	4	ABG07902	ABG07902 Novel hum
26	741	45.5	140	5	ABR52404	ABR52404 Peptide r
27	740	45.5	117	5	AAU79159	AAU79159 Human dua
28	716	44.0	140	5	ABR52405	ABR52405 Peptide r
29	715.5	44.0	579	7	ADBE08458	ADBE08458 Novel pro
30	654.5	40.2	170	4	AAE66436	AAE66436 Human MAP
31	654.5	40.2	170	4	AAE06780	AAE06780 Human pho
32	476	29.3	482	4	AAE73225	AAE73225 Human pho
33	469	28.8	444	5	AAO20515	AAO20515 Protein o
34	469	28.8	482	3	AAI18655	AAI18655 A human r
35	469	28.8	482	3	AAI29641	AAI29641 Human dua
36	469	28.8	482	5	ABR52384	ABR52384 Protein r
37	469	28.8	482	5	ABJ05600	ABJ05600 Breast ca
38	467.5	28.7	381	5	ABR52383	ABR52383 Protein r
39	467.5	28.7	381	6	ABP96803	ABP96803 Human COP
40	462.5	28.4	381	6	ABP57087	ABP57087 Mouse MKP
41	461.5	28.4	381	7	ADD48300	ADD48300 Rat Prote
42	461.5	28.4	381	7	ADG62625	ADG62625 Rat Prote
43	440.5	27.1	368	5	AAE14240	AAE14240 Human pro
44	440.5	27.1	419	6	ABR55026	ABR55026 Human dua
45	428	26.3	394	4	AAE67627	AAE67627 Amino aci

ALIGNMENTS

RESULT 1	ABR52381	ABR52381 standard; protein; 665 AA.
ID	ABR52381;	
AC	ABR52381;	
XX	19-JUN-2003	(first entry)
DE	Protein relating to the invention SEQ ID NO: 109.	
XX		
XX	anti-proliferative; hepatotropic; nephrotropic; antiarthritic;	
KW	anti-proliferative; cardiact; cytosolic; gene therapy; liver disease;	
KW	proliferative disorder; renal failure; cardiovascular disorder;	
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;	
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200257460-A2.	
XX		
PD	25-JUL-2002.	
XX		
PF	20-DEC-2001; 2001WO-US050459.	
XX		
PR	20-DEC-2000; 2000US-025668P.	
PR	30-MAR-2001; 2001US-028018P.	
PR	01-MAY-2001; 2001US-0287735P.	
PR	05-JUN-2001; 2001US-0295848P.	
PR	25-JUN-2001; 2001US-0300465P.	
XX		
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.	
XX		
PI	Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;	
PI	Stemers N, Bol D, Schieven G, Finger J, Todderud CO, Basolino D;	

PI Krystek S, Mcatee P, Suchard S, Banas D,
 XX WPI; 2002-599721/64.
 DR N-PSDB; ACC60559.
 XX
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 XX
 PS Claim 5; Fig 12; 801pp; English.

CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatotropic, nephrotropic,
 CC antiarthritic, antipsoriatic, cardiant and cytostatic activity. The
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 XX
 SO Sequence 665 AA;

Alignment Scores:

Pred. No.: 1 18e-170
 Score: 1564.00 Length: 665
 Percent Similarity: 100.00% Matches: 302
 Best Local Similarity: 100.00% Conservative: 0
 Query Match: 96.13% Mismatches: 0
 Gaps: 0
 Indels: 0

US-10-029-345a-108_COPY_538_1443 (1-906) x ABR52381 (1-665)

QY 1 ATGGCCCATGAGATGTTGAACTCAAAATGTTACTGAGAGGTGGCTGCTGCGAA 60
 DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
 QY 61 AGTGAACGGAAGAAAGGCTGCTAAATGATAGCCGCCATTTGGTGAATACATACATCC 120
 DB 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 QY 121 CACATTTGGAGCCATTAAATATCAACTGCTCCAGCTTATGAGAGCGAAGTTGCAACG 180
 DB 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerIleLeuMetIleValGlnGlnGln 60
 QY 181 GACAAAGTGTATATACAGAGCTCATCCAGCTTACGAAACATTAAGTTGACATTGAT 240
 DB 61 AspIleValIleuIleThrGlnLeuIleGlnHisSerIleAlaIleValIleAsp 80
 QY 241 TGCAGTCAGAGGTTGATGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTTCA 300
 DB 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIleValIle 100
 QY 301 GACGTTTTCACGTACTCTGAGTAACTGGAAGAGCTTCACTCTGTTCACTG 360
 DB 101 AspCysPheLeuThrValIleLeuGlyIleValLeuGlnIleValIleValIleValIle 120
 QY 361 CTTCAGAGTGGGTTTGCAGATTCCTGCTGTTGTTTCCCTGCGCTCTGTAAGAAATCC 420
 DB 121 LeuAlaGlyIlePheAlaGlnPheSerArgCysPheProGlyIleuCysGlnIleValIle 140
 QY 421 ACTTACGCTTACCTGATTTCTCAGCCCTTGCTTACCTGTTGCAACAATGGGCAAC 480
 DB 141 ThrIleValProThrCysIleSerGlnProCysIleProValAlaAsnIleGlyProThr 160
 QY 481 CGAATTTCCCAATCTTATCTTGCTCCGACGAGATGTCCTCAACAAGAGCTGATA 540
 DB 161 ArgIleLeuProAsnLeuIleValIleGlyCysGlnIleArgPheValIleAsnIleValIle 180

QY 541 CAGCAGATGAGATTTGTTATGTGTTAAATGACAGCTATACCTGCAAGGCTGACTTT 600
 DB 181 GlnHisAsnIleIleIleIleValLeuAsnIleSerIleThrCysProIleProAspPhe 200
 QY 601 ATCCCGAGTCTCATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIleLeu 220
 QY 661 CCGTGGTGGACAAATCTGATGATTTCAATGAGAAAGCAAGCTTCAATGATGATGAT 720
 DB 221 ProIlePheAspIleValIleValIleValIleValIleValIleValIleValIleVal 240
 QY 721 CTATGCACTGTTTACGTGGATCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 780
 DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerIleThrIleAlaIleValIleMet 260
 QY 781 AAGAGATGAGCAGCTCTTATGATGAGCTTACAGATTTGTGAAAGAAAGAAAGCTACT 840
 DB 261 LysArgPheCysPheSerIleuAspGlnAlaIleValIleValIleValIleValIleVal 280
 QY 841 ATATCTCAAACTTCAATTTCTGAGCAACTCTGAGCTATGAGAGAAAGATTAAGAAC 900
 DB 281 IleSerProAsnPheAsnPheLeuGlyIleLeuLeuAspIleValIleValIleValIle 300
 QY 901 CAGACT 906
 DB 301 GlnThr 302

RESULT 2
 ABR52407
 ID ABR52407 standard; protein; 665 AA.
 XX
 AC ABR52407;
 XX
 XX 19-JUN-2003 (first entry)
 DT
 XX
 DE Protein relating to the invention SEQ ID NO: 148.
 XX
 KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;
 KW antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
 KW proliferative disorder; renal failure; cardiovascular disorder;
 KW immunological disorder; arthritis; psoriasis; congenital heart defect;
 KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
 OS Homo sapiens.
 XX
 XX
 PN WO200257460-A2.
 XX
 PD 25-JUN-2002.
 XX
 PF 20-DEC-2001; 2001WO-US050459.
 XX
 PR 20-DEC-2000; 2000US-025686BP.
 PR 30-MAR-2001; 2001US-0280186P.
 PR 01-MAY-2001; 2001US-0287735P.
 PR 05-JUN-2001; 2001US-0295848P.
 PR 25-JUN-2001; 2001US-0300465P.
 XX
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA
 PI Jackson DG, Feder J, Nelson T, Mintler G, Ramanathan C, Lee L,
 PI Siemers N, Bol D, Schieven G, Finger J, Todderd CG, Basolino D,
 PI Krystek S, Mcatee P, Suchard S, Banas D,
 DR WPI; 2002-599721/64.
 DR N-PSDB; ACC605572.
 XX
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 PS Disclosure; Fig 19; 801pp; English.

XX The invention relates to a novel isolated nucleic acid comprising a CC polynucleotide having a nucleotide sequence selected from 40 CC polynucleotides fully defined in the specification. The polynucleotide of CC the invention has antiproliferative, hepatotropic, nephrotropic, CC antiarthritic, antipsoriatic, cardant, and cyclostatic activity. The CC polynucleotide may have a use in gene therapy. A polynucleotide or CC polypeptide of the invention is useful for preventing, treating or CC ameliorating a medical condition, e.g. a proliferative disorder. They are CC also useful for treating e.g. liver disease, renal failure, immunological CC disorders including arthritis and psoriasis, cardiovascular disorders CC such as congenital heart defects and congestive heart failure, and CC cancer. A method of the invention is useful for diagnosing a pathological CC condition or susceptibility to a condition in a subject. The present CC sequence is used in the exemplification of the invention

SQ Sequence 665 AA;

Alignment Scores:	
Pred. No.:	1,18e-170
Score:	1564.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	96.13%
DB:	5
Length:	665
Matches:	3020
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) X ABR52407 (1-665)

[illegible]

Db	221	ProtripeuaplyserserValaspPheIleGluysalAllysaseramIlyCysVal	240
Qy	721	CTATGCACTGTTTAGCTGGAGATCCCGGCTCCGGCACATCGCATCGCTACATCATG	780
Db	241	LeuValIhscysleuAlagIyIleserxrgseralathrlleAllelelathyllleMet	260
Qy	781	AAGAGATGACATGCTCTTTAGATGAGACCTTACAGATTGTGAAAGAAAAGACTTACT	840
Db	261	LysrIgwetIasPmetSerleuaspIulAlatyrArgPheValIlyglIuysArgProthr	280
Qy	841	ATATCTCCAAACTTCATTTTCTGGGCCCACTCCTGACATATAGAAAGAAAGATTAAAGAC	900
Db	281	IleserProanPheanPheleuGllyGlnleuenuapTyGIuIyysrlleuysan	300
Qy	901	CAGACT 906	
Db	301	Glnthr 302	

RESULT 3
AAE04834
ID AAE04834 standard; protein; 665 AA.

AC AAE04834;

DT 10-SEP-2001 (first entry)

DE Human GGP002 phosphatase polypeptide.

KM Human, PNP002 phosphatase polypeptide; phosphatase-related disease;
KM immune-related disorder; ocular disease, organ transplant rejection;
KM infection, diabetes, pain, sexual dysfunction, Alzheimer's disease;
KM metabolic disorder, hemangioma, hemangioma, cancer, mood disorder, cardiac;
KM Parkinson's disease, multiple sclerosis; amyotrophic lateral sclerosis;
KM cardiovascular disease, brain; neuronal-associated disease, dyskinesia;
KM attention disorder, cognition disorder, psychotic disorder, cytotoxic;
KM neurological disorder, virulence, nocturnal; cerebroprotective, therapy;
KM neuroprotective, antibacterial, vulnarity, tranquility, antineoplastic;
KM hypotensive, immunosuppressive, antipsychotic, analgesic, hypertensive;
KM antifungal; dual specificity phosphatase, DSP, MAP kinase phosphatase;
KM MKP, migraine; chromosome 12p11.1-1912.1.

OS Homo sapiens.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

```
FT /label= Catalytic_domain
```

```
/label= Phosphatase_domain
```

PN WO200146394-A2.

PD 28-JUN-2001.

PF 21-DEC-2000; 2000WO-US034736.

PR 21-DEC-1999; 99US-0173255P.

PR 25-JAN-2000; 2000US-0178078P.

XX

XX

PI Hill RJ, Flanagan P;

DR WPI; 2001-418058/44.

XX

PT diseases and disorders, cardiovascular disease, brain or neuronal-

XX

PS Claim 7; Fig 2; 186pp; English.

xx

CC The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorders, including cancers of tissues, cancers of
CC haematopoietic origin, diseases of central and peripheral nervous system
CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
CC lateral sclerosis, viral infections, infections caused by prions,
CC bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual
CC dysfunction, mood disorders, attention disorders, cognitive disorders,
CC hypertension, hyperextension, psychotic disorders, neurological disorders,
CC dykinesias and organ transplant rejection. The present amino acid
CC sequence is human SGP002 phosphatase polypeptide. This sequence is
CC classified as dual specificity phosphatase (DSP) and MAP kinase
CC phosphatase (MMP). SGP002 gene maps to chromosomal position 12p11.1-p12.1
CC
CC
CC Sequence 665 AA;
SQ

Alignment Scores:	
Pred. No.:	2,856-169
Score:	1562.00
Percent Similarity:	99.674
Best Local Similarity:	99.344
Query Match:	95.394
DB:	4
Length:	665
Matches:	300
Conservative:	1
Mismatches:	0
Indels:	0

----- (1-906) X AAE04834 (1-665)

1	ATGCCCATGAGTATGATTGGAACTCAAAATTGTACTGAGAGGTGGTGGCTGCTGGGAA	60
2	1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln	20
3	61 AGTGAACGGAAAAAATGCTGCTAAATTGATAGCCGGCCATTGTGGAAATATACATCATCC	120
4	21 SerGlyThrGlnLysValLeuLeuIleAspSerHisGlyProPheValGlnLysThrSer	40
5	121 CACATTTTGGAAAGCCATTATATATCACTGCTCCCAAGCTTATGTAAGCGAAGTTGGCAAC	180
6	41 HisIleLeuGlnAlaIleAsnIleAsnLysSerLysLeuMetLysArgLysGlnGln	60
7	181 GACAAAGCTTAAATTAACAGAGCTTCATCCAGACATTTCAGAGAAACATTAAGTTGACATTGAT	240
8	61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp	80
9	241 TGCAGTCAAGAGGTTGTAGTTTACATCAAAAGCTCCAAAGATGTTGGCTCTCTCTCTTA	300
10	81 CysSerGlnLysValValAlaLysAspGlnSerSerGlnAspValAlaSerLysSerSer	100
11	301 GACGTTTCTTCACTGACTTCTGGGGTAACTGAGAAAGAGCTTCAACTCTGTTCACTG	360
12	101 AspCysPheLeuThrValLeuLeuGlnLysLysLeuGlnLysSerPheAsnSerValHisLeu	120
13	361 CTTCAGAGGGGTTTCTGAGTCTCTGTTGTTCCCTGAGCTCTGTGAAGAAATCC	420
14	121 LeuAlaGlyAlaPheAlaGlnPheSerArgCysPheProGlyLeuGlySerLysSer	140
15	421 ACTCTAGTCCCTACCTGCAATTTCTCAGGCTTGCTTACCTGTTGCCAAATTTGGGCAAC	480
16	141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
17	481 CGAATCTTCCCAATCTTATCTTGCTGGCTGCGACGAGATGTCTCAACAAAGAGCTGATA	540
18	161 ArgIleLeuProHisLeuIleLysGlyCysGlnArgAspValLeuAsnLysGlnLeuMet	180
19	541 CAGCGAAATGGAGATTGGTTATGTGTTAAATGCGACAGTATACCTGTCCAAAGCTGACTTT	600
20	181 GlnGlnAsnGlnIleGlyLysValLeuAsnAlaSerAsnThrCysProLysProAspPhe	200
21	601 ATCCCGAGTCTCATTTCTCGCGCTGCGCTGTGAATGACAGACTTTTGTGAGAAATTTTG	660

Db	201	IIePFGIuSerHsIaPheLeuAaGValPrlvaAlaAspSerPheCysGIuLyIleLeu	22
QY	661	CGGTGTTGGCAAAATGATGATTCATTGAGAAAAGCAAAAGCTCCCAATGATGT	72
Db	221	ProTrpLeuAspLysSerValaAspHeiIeGIuLysAlaLysAlaSerHndIcysVal	240
QY	721	CTAGTGACATGTTTAGCTGGATCTCCGCTCGGCACATCGGATCGCTCATCATG	780
Db	241	LeuValHscysLeuAlaGIyIleSerAgsSerAlaThrIleAlaIleAlaIylIleMet	260
QY	781	AAGAGATGACATGTCCTTTAGATGAAAGCTTACAATTTGGAAAAGAAAAGACTACT	840
Db	261	LysAlrMetAspMetSerLeuAspGIuAlaIyrAlaGpHeuAllySGIuLysGProThr	280
QY	841	ATATCTCCAAACTTCATTTCTGGGCCAATCTCTGACATAGAGAAGAAGATTAAAC	900
Db	281	IleSerProAsnPhaAsnPhaLeuGIyGIuLeuLeuAspTyrGIuLysIleLysaen	300
QY	901	CAGACT 906	
Db	301	GIuThr 302	

RESULT 4	
AAU09016	
ID	AAU09016 standard; protein; 665 AA.
XX	

AC AAU09016;
XX

18-DEC-2001 (first entry)

specificity phosphatase 21117.

immunogen.
immunogen.
acute myeloid leukaemia;
cellular proliferative; differentiative disorder; leukaemia;
liver disorder; erythroid associated disorder; haemolytic anaemia;
rheumatoid arthritis; multiple sclerosis; Crohn's disease;
neuropathic disorder; autoimmune disorder; diabetes mellitus;
hematopoietic disorder; autoimmune disorder; hepatotropic; cytostatic;
immunogen.

Homo sapiens.

Domain	Location/Qualifiers
11	131

Domain	/label= Rhodanese-like domain
158..297	

```

/label= catalytic domain
/note= "Dual specificity

```

```
/label= Tyrosine specific protein phosphatase
```

WO200173059-A2.

04-OCT-2001

23-MAR-2001; 2001WO-US009477.

2000US-0191858P.

ALLENBURY PHARM INC.

WDT
1000

N-PSDB; AAS14639.

New human dual SF
of disease and

Claim 9: Fig 1: 143m. E-1'1'

Two novel human dual specificity phosphatases

CC designated 2117 and 38692, the nucleic acids encoding them (including
 CC fragments, allelic variants, their complements or nucleic acids that
 CC hybridize to them) and antibodies raised against the proteins. The
 CC antibody is useful for detecting the presence of the polypeptide, and the
 CC nucleic acid fragments are useful for detecting the presence of the
 CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
 CC antisense sequences) are useful for modulating the activity or expression
 CC of the polypeptide/nucleic acid, useful for the treatment of e.g.
 CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
 CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
 CC listed in the specification) liver disorders, erythroid associated
 CC disorders (e.g. haemolytic anaemia) cellular proliferative or
 CC differentiative disorders, leukaemias (e.g. acute myeloid leukaemia),
 CC carcinomas, sarcomas, metastatic cancers. Agents modulating 2117 and
 CC 38692 are also useful for modulating the proliferation, survival,
 CC migration or differentiation of a 38692 or 2117-expressing cell. The
 CC polypeptide and nucleic acids are useful for identifying modulating
 CC agents. The present sequence represents the dual specificity phosphatase
 CC 2117

XX Sequence 665 AA,

Alignment Scores:

Pred. No.: 2,85e-169 Length: 665
 Score: 1552.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 0
 Query Match: 95.39% Indels: 1
 DB: 4 Gaps: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x AAU09016 (1-665)

QY 1 ATGGCCCATGAGATGATGGAAGTCTTACTGAGAGGTTGGTGGCTCTGCTGGA 60
 Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuGln 20
 QY 61 AGTGAACGGAAGAAGTGTGTAATTGATGACCGGCACTTTGTGGAATCAATACATCC 120
 Db 21 SerGlyThrGlnIleValLeuLeuIleAspSerThrProPheValGlnTyrAsnThrSer 40
 QY 121 CACATTTGGAGCCATTAAATCACTGCTCCAGCTTATGAGAGGAGTTGCAACAG 180
 Db 41 HisIleLeuGlnIleAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 60
 QY 181 GAAAGAGTGTATTCAGAGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 240
 Db 61 AspIleValLeuIleThrGlnIleGlnHisSerIleValHisIleValAspIleAsp 80
 QY 241 TGCAGTCAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 Db 81 CysSerIleIleValIleValIleValIleValIleValIleValIleValIleVal 100
 QY 301 GACTGTTTCTCAGTCTCTGAGTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 Db 101 AspCysPheLeuThrValIleLeuGlnIleValIleValIleValIleValIleVal 120
 QY 361 CTTCGAGAGGTTGTTGCTGAGTCTCTGCTGTTTCTCTGCTGCTGCTGCTGCTGCT 420
 Db 121 LeuAlaGlyIleGlyPheAlaGlnPheSerIleValIleValIleValIleValIleVal 140
 QY 421 ACTCTAGTCCCTCAGCAGTCTTCTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATCTTCCCAATCTTATCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 161 ArgIleLeuProAsnLeuIleValIleValIleValIleValIleValIleValIleVal 180
 QY 541 CAGCAGAAATGGATGGTATGCTTAAATGCAAGCAGTATCTGCTGCAAGCTGACTTT 600
 Db 181 GlnGlnAsnIleIleGlyIleValIleValIleValIleValIleValIleValIleVal 200
 QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

Db 201 IleProGlnSerHisPheLeuArgValProValAlaAspSerPheCysGlnIleLeu 220
 QY 661 CGCTGTTGGACAAATCAGTACTGATTTTCATTGAGAAAGCAAGCCTTCAATGATGTT 720
 Db 221 ProTrpLeuAspIleSerValAspPheIleGlnIleValAlaIleValIleValIleVal 240
 QY 721 CTGAGCAGCTGTTTACTGGGATCTCCGCTCGCCACCATGCTGCTGCTGCTGCTGCTGCT 780
 Db 241 LeuValHisCysLeuAlaGlnIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
 QY 781 AACGAGATGACATGCTTCTTGAATGAGCTTACAGATTTGTGAAAGAAAAGACCTACT 840
 Db 261 LysArgMetAspMetSerIleuAspGlnIleValIleValIleValIleValIleValIleVal 280
 QY 841 ATATCTTCAACTTCAATTTTCTGGCCCAACTCTGAGCTATGAGAAAGAAAGAAC 900
 Db 281 IleSerProAsnPheAsnPheLeuGlnIleLeuLeuAspTyrGlnIleValIleValIleVal 300
 QY 901 CAGACT 906
 Db 301 GlnThr 302

RESULT 5
 AAU79156
 ID AAU79156 standard; protein; 665 AA.
 XX AAU79156;
 AC
 XX
 DT 02-JUL-2002 (first entry)
 DE Human dual-specificity phosphatase-3 (DSP-16) protein.
 XX
 XX Human, dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; chromosome 12p.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 242..251
 PN WO200226997-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 25-SEP-2001; 2001MO-US030124.
 XX
 PR 26-SEP-2000; 2000US-0235487P.
 XX
 PA (CEPR-) CEPTYR INC.
 XX
 PI Lucie RM, Wei B;
 XX
 XX MPI; 2002-315802/35.
 DR N-PSDB; ABK47596.
 XX
 PT New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 PS
 PS Claim 1; Fig 2; 87pp; English.
 XX
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival

CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC gene located on chromosome 12p
 CC
 SQ Sequence 665 AA;

Alignment Scores:

Pred. No.:	2.85e-169	Length:	665
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	0
Query Match:	5	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) x AAU79156 (1-665)

```

QY 1 ATGGCCCATGAGTGAATGGAATCAATTTGTTACTGAGAGTGGCTGCTGCGAA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTGCTGCTAATTTGATGCGCGCATTTGGAATACATACATCC 120
DB 21 SerGlyThrGlnIleValLeuLeuIleAspSerArgProPheValGluIleThrSer 40
QY 121 CACATTTTGGAAAGCCATTAATATCACTGCTCCAAAGCTTATGAAGGAGAGTTGCAACG 180
DB 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerIleuMetIleArgIleuGln 60
QY 181 GACAAAGTGAATTAATCAAGAGTTCACAGATTCACAGAAACATTAAGTTGACATGAT 240
DB 61 AspIleValIleuIleThrGlnIleuIleGlnHisSerAlaIleHisIleValAsp 80
QY 241 TGCAGTAGAAGGTTGATGTTACGATCAAGAGCTCCAGATGTTGCTCTCTCTTCA 300
DB 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIleVal 100
QY 301 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAGAGAGTTCACTGTTACCTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlnIleValIleValIleValIleValIleVal 120
QY 361 CTGCAAGGTTGGTTTCTGAGTCTCTGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 121 LeuAlaGlyIleValPheAlaGlnPheSerArgCysPheProGlyLeuGlyIleVal 140
QY 421 ACTTAGTCCCTACCGCATTTCTAGGCTTCTTACTGTTGCGCAACATTTGGCGCAC 480
DB 141 ThrIleValProThrCysIleSerGlnProCysIleuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTATCTTGGCTGCGCAGAGATGCTCTCAACAAGAGCTGAGA 540
DB 161 ArgIleLeuProAsnIleuIleuIleGlyCysGlnArgPheValLeuAsnIleGlnMet 180
QY 541 CAGCAAGATGGGATTTGTTATGTTAAATGCGAGCTATCTGTCGCAAGCTTCACTT 600
DB 181 GlnGlnAsnGlnIleGlyIleValIleAsnIleSerAsnIleThrCysProIleProAspPhe 200
QY 601 ATCCCGAGTCTGATTTCTGCGGTGCTGCTGATGAGATGAGCTTTTGTGAGAAATTTTG 660
DB 201 IleProGlnIleSerHisPheLeuArgValIleProValAsnAspSerPheCysGlnIleValIleu 220
QY 661 CCGGTGTTGACAATATGATGATTTCAATGAGAAAGAAAGCTTCCATATGATGTT 720
DB 221 ProIlePheAspIleSerValIlePheIleGlyIleValIleValIleValIleValIleVal 240
QY 721 CTAGTGACATGTTTACTGATGATCTCCGCTCCGACCATGCGATATGCTTACATGATG 780
DB 780

```

```

DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
QY 781 AAGAGATGAGACATGCTTTTATATGATACCTTACAGATTTTGAAAGAAAGAAAGCTACT 840
DB 261 LysArgMetAspPheSerIleuAspIleValIleValIleValIleValIleValIleVal 280
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTCGATGATGAGAGAAAGATTAAGAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIleGlnIleValIleValIleVal 300
QY 901 CAGACT 906
DB 301 GlnThr 302

```

RESULT 6

AAU09946 ID AAU09946 standard; protein; 665 AA.

AAU09946;

18-JUN-2002 (first entry)

Protein sequence of human (dual specificity phosphatase) DSP-10.

Human; dual specificity phosphatase; DSP-10; cancer; epilepsy; stroke;
 neuronal degeneration syndrome; Alzheimer's disease; depression;
 schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
 osteoporosis; diabetes.

Homo sapiens.

WO200177340-A1.

18-OCT-2001.

06-APR-2001; 2001WO-BP003966.

10-APR-2000; 2000EP-00107143.

(MERCK) MERCK PATENT GMBH.

Duecker K;

WPI; 2002-010917/01.

N-PSDB; AAS15768.

Novel dual specificity phosphatase polypeptides useful for treating
 cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
 disease, depression, schizophrenia, asthma and immune disorders.

Claim 2; Page 37-39; 43pp; English.

The present invention relates to a new isolated dual specificity
 phosphatase (DSP10) polypeptide, comprising a 665 residue amino acid
 sequence that is fully defined in the specification. The invention also
 provides a sequence encoded by a 3059 nucleotide sequence fully defined
 in the specification, and a sequence having at least 95 % identity to the
 polypeptide, or fragments or variants of DSP-10. The invention is useful
 for treating cancer e.g. leukemia, colon carcinoma, lung cancer,
 prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
 stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
 schizophrenia, cardiac myopathies, asthma, immune disorders,
 inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
 osteoporosis, diabetes and diabetes associated diseases. The molecules of
 the invention are also useful as vaccines for inducing immunological
 response in a mammal, in disease diagnosis and in assays for screening
 agonistic or antagonistic compounds. Other uses of the invention include
 identifying membrane bound or soluble receptors, as a diagnostic reagent,
 in chromosome localisation studies, and as a valuable tool in tissue
 expression studies. The present sequence represents the dual specificity
 phosphatase, DSP-10, protein of the invention

Sequence 665 AA;

CC treating or preventing disorders associated with aberrant expression of
 CC protein phosphatases (PP), particularly immune system disorders e.g.
 CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
 CC asthma or Crohn's disease, neurological disorders e.g. epilepsy,
 CC Huntington's disease, dementia or Parkinson's disease, developmental
 CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
 CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
 CC or sarcoma. The present amino acid sequence represents human protein
 CC phosphatase 7 (PP7) which is one of several human protein phosphatases
 CC (AAU05783-AAU05792) of the invention
 XX
 SQ Sequence 665 AA;

Alignment Scores:
 Pred. No.: 2,85e-169 Length: 665
 Score: 1552.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 1
 Query Match: 95.39% Indels: 0
 DB: 5 Gaps: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x AAU05789 (1-665)

QY 1 ATGGCCCATGATGATGGAAGTCTCAATTTGTTACGAGAGTTGGTCTGCTGAA 60
 DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
 QY 61 AGTGAACGGAAAAAGTGGCTGAATGATGATGAGCGGCAATTTGGAAATCAATACATCC 120
 DB 21 SerGlyThrGluValLeuValLeuLeuIleAspSerArgProPheValGluIuYrAsnThrSer 40
 QY 121 CACATTTGGAGCCCATTAATATCACTGCTCCAGCTTATGAAGCGAAGTTGGACACG 180
 DB 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleuMetIleValArgLeuGlnGln 60
 QY 181 GACAAAGTGTATATACAGAGCTCATCCAGATTCAGCCAAACATTAAGTTGACATTGAT 240
 DB 61 AspIleValIleuIleThrGluIleIleGlnHisSerIleAlaIleHisIleValAspIleAsp 80
 QY 241 TGCAGTCGAGAGGTTGATGTTACGATCAAGCTCCCAAGATGTTGCCCTCTCTCTTCA 300
 DB 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIleValIleVal 360
 QY 301 GACTGTTTCTCAGCTGACTTCTGGGTTAACTGGAAGAAAGCTTCACTGTTACCTG 360
 DB 101 AspCysPheLeuThrValIleuLeuGlyIleuGlnIleuGlnIleuGlnIleuGlnIleu 120
 QY 361 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTTTCCTGGCCTCTGTGAAGAAATCC 420
 DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlyIleuSer 140
 QY 421 ACTTGTGCTCTACCTGATTTCTCAGCTTCTGCTTACCTGTTGGCAACATTTGGCCCAACC 480
 DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATCTTCCCAATCTTATCTTGCTGCTGCGCAGAGATGCTCTCAACAAGAGCTGATA 540
 DB 161 ArgIleLeuProAsnIleuIleuGlyCysGlnArgPheValIleuAsnIleuGlnIleuGln 180
 QY 541 CAGCAGATGGAGTTGTTATGTGTAAAGCAGCATATCTGTCGAAGCTTCACTTT 600
 DB 181 GlnGlnIleuGlnIleuGlyIleuGlyIleuValIleuAsnIleuSerIleuThrCysProIleuProAspPhe 200
 QY 601 ATCCCGAGTCTCATTTCTGCGGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 660
 DB 201 IleProGlnIleuSerIleuPheLeuArgValProValAsnAspSerPheCysGluIleuIleu 220
 QY 661 CCGGCTTGGACAAATCATGATGATTTGATGAGAAAGAAAGCTCCAAATGATGTTG 720
 DB 221 ProIlePheAspIleuSerValIleuPheIleuGlyIleuGlyIleuGlyIleuGlyIleuGly 740
 QY 721 CTAGTGACCTGTTTACCTGGAGTCCCGCTCCGACCATGCTATGCTTACATATG 780

DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260
 QY 761 AAGAGATGACATGTTGTTAGATGACCTTACAGATTTGGTGAAGAAAAAGACTTACT 840
 DB 261 LysArgMetAspMetSerIleuAspGluIleValIleValIleValIleValIleValIleVal 280
 QY 841 ATATCTCCAACTTCAATTTCTGGGCCAAGCTCTGCACTATGAGAGAAAGATTGAAC 900
 DB 281 IleSerProAsnPheAsnPheLeuGlyGlnIleuLeuAspTyrGlnIleuIleuIleuAsn 300
 QY 901 CAGACT 906
 DB 301 GlnThr 302

RESULT 8
 ID ABB97946
 ID ABB97946 standard; protein; 665 AA.
 AC ABB97946;
 XX
 DT 06-SEP-2002 (first entry)
 XX
 DE Human protein sequence #13.
 XX
 KW Human; brain; tonsil; hippocampus; foetal brain; diagnosis.
 OS Homo sapiens.
 XX
 PN W0200252005-A1.
 XX
 PD 04-JUL-2002.
 XX
 PE 20-DEC-2001; 2001WO-JP011217.
 XX
 PR 22-DEC-2000; 2000JP-00389742.
 XX
 PA (KAZU-) KAZUSA DNA RES. INST. FOUND.
 PA (CELE-) CELESTAR LEXICO-SCI LTD.
 PI Ohara O, Nagase T, Nakajima D;
 XX
 DR WPI, 2002-500762/53.
 DR N-PSDB; ABB97946.
 XX

Genes and their expression products cloned from human cDNA libraries for treatment and diagnosis of diseases associated with their expression.
 Claim 1(a); Page 112-116; 238pp; Japanese.

The invention relates to DNA encoding polypeptides directly cloned from cDNA libraries originating in adult whole brain, human tonsil, human adult hippocampus and human foetal brain. Polypeptides and polynucleotides of the invention may be used in the investigation of differential expression of the DNA sequences in normal subjects and disease patients. They may also be used in the production of antibodies and oligonucleotide probes and DNA chips for diagnosis and identification of drugs for treatment of diseases with which the DNA sequences are associated. The sequences given in records ABB97934-ABB97964 represent human proteins of the invention

SQ Sequence 665 AA;

Alignment Scores:
 Pred. No.: 2,85e-169 Length: 665
 Score: 1552.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 1
 Query Match: 95.39% Indels: 0
 DB: 5 Gaps: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x ABB97946 (1-665)

QY 1 ATGGCCCATGATGATGGAAGTCTCAATTTGTTACGAGAGTTGGTCTGCTGAA 60

DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
QY 61 AGTGGAAAGGAAAGAGTGGCTGCTAATTGATAGCCGGCATTGGTAATACATACATCC 120
DB 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
QY 121 CACATTTTGGAGCCATTATATCACTGCTCCAGCTTGTGAAGGAGGTTGCAACG 180
DB 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerLeuMetLysArgArgLeuGlnGln 60
QY 181 GACAAAGTTAAATTAACAGAGCTATCCAGCATTCAGCGAAATAGATTGACATTGAT 240
DB 61 AspLysValLeuIleThrGluLeuIleGlnHisSerLysHisValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATTACGATCAAAAGCTCCCAAGATGGCTCTCTCTTCA 300
DB 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTCACTGACTTCTGGGTAACCTGGAGAGAGCTTCACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGTGGGTTTGGCTGAGTTCTCTCGTTGTTCCCTGGCCTCTGTGAAGGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyLysSer 140
QY 421 ACTGTAGTCCCTACCTGACATTTCTCAGCCTTGTCTTACCTGTTGCCAATTTGGCCCAACC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTTATCTTGCTGCGACGAGATGCTTCCAAACAGAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuValLysGlyCysGlnArgAspValLeuAsnLysGluLeuMet 180
QY 541 CACAGAGATGGATTGGTTATGTTAAATGACAGCTATACCTGTCCAAAGCCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGCTGCTGCTGTAATGACAGCTTTTGTGAAAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
QY 661 CCGTGTGTGACAAATCAGTAGATTTCATGAGAAAGCAAAAGCCCAATGATGTT 720
DB 221 ProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTAGTCACTGTTAGCTGGGATCTCCGCTCCGCCACATCGCTATCGCTACATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerLisThrIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGATGAGACATGTTTGAATGAGCTTACAGATTGTGAAGAAAAAGAAAGCTTACT 840
DB 261 LysArgMetSerMetSerLeuAspGlnAlaTyrArgPheValLysGluLysArgProThr 280
QY 841 ATATCTCCAAACTTCATTTTTCGGGCGCACTCTGAGCTATGAGAAAGAAAGATTAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLysAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302

KM Human; dual specificity phosphatase 21117; erythroid-related disorder;
KW haemopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
KM erythrocytosis; liver-related disorder; cancer.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 11..131
FT 1/label= Rhodanese_like_domain
FT Region. 21..24
FT /note= "Casein kinase II phosphorylation site"
FT Region 91..94
FT /note= "Casein kinase II phosphorylation site"
FT Domain 158..297
FT /note= "Dual specificity phosphatase catalytic domain"
FT Region 214..217
FT /note= "Casein kinase II phosphorylation site"
FT 242..254
FT /note= "Tyrosine specific protein phosphatase active site"
FT Region 242..254
FT /note= "C-X5-R motif"
FT Region 266..269
FT /note= "Casein kinase II phosphorylation site"
FT Region 369..372
FT /note= "Casein kinase II phosphorylation site"
FT Region 421..424
FT /note= "Casein kinase II phosphorylation site"
FT Region 434..437
FT /note= "Casein kinase II phosphorylation site"
FT Region 458..461
FT /note= "Casein kinase II phosphorylation site"
FT Region 508..511
FT /note= "Casein kinase II phosphorylation site"
FT Region 589..592
FT /note= "Casein kinase II phosphorylation site"
FT Region 612..615
FT /note= "Casein kinase II phosphorylation site"
FT Region 617..620
FT /note= "Casein kinase II phosphorylation site"
FT Region 642..645
FT /note= "Casein kinase II phosphorylation site"
PN US002034807-A1.
XX
XX 21-MAR-2002.
XX
XX 23-MAR-2001; 2001US-00816494.
XX
XX 24-MAR-2000; 2000US-0191858P.
XX
XX (MEYE/) MEYERS R A.
XX
XX Meyers RA;
XX
XX MPI; 2002-351088/38.
XX N-PSDB; ABK49402.
XX
XX New nucleic acids, designated 38692 and 21117, encoding dual specificity
PT phosphatases for treating cell proliferation and differentiation
PT disorders including hematopoietic and erythroid-related disorders and
PT cancers.
XX
XX Claim 8; Fig 1; 76pp; English.
XX
XX The present invention relates to new nucleic acids designated 38692 and
XX 21117 encoding dual specificity phosphatase family members. The nucleic
XX acid, polypeptide encoded by it, and antibody specific for the
XX polypeptide may be used to diagnose and treat haemopoietic-related
XX disorders such as leukaemias and autoimmune diseases, erythroid-related
XX disorders such as anaemias and erythrocytosis, liver-related disorders,
XX and cancers, particularly of the breast, colon, adipose, prostate and
XX lung. The present amino acid sequence represents the human dual

CC specificity phosphatase 21117 protein of the invention, as described
 CC above
 XX
 SQ Sequence 665 AA;

Alignment Scores:

Pred. No.: 2,85e-169 Length: 665
 Score: 1552.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 1
 Query Match: 95.39% Indels: 0
 DB: 5 Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x AAU79929 (1-665)

```

QY 1 ATGGCCCATGAGATGATTTGGAAGTCAAAATGTTACTGAGAGGTGGTGGCTCTGCTGGA 60
    |||||
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 60
QY 61 AGTGAACGGAAGAAAGTGGCTTAATGATAGCCGGCCATTGTGGAATACAAATCATCC 120
    |||||
Db 21 SerGlyThrGlnValLeuLeuIleAspSerArgProPheValGluIlyAsnThrSer 40
QY 121 CACATTTTGAAGCATTAATATCACTGCTCCAGCTTATGAAAGGAGGTTGCAAG 180
    |||||
Db 41 HisIleuGlnAlaIleAsnIleAsnCysSerIlySleuMetIlyArgValGlnGln 60
QY 181 GACAAAGTTAATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 240
    |||||
Db 61 AspIlyValLeuIleThrGluLeuIleGlnHisSerAlaIlySlyValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGTTACTGATTCGATCAAAAGCTCCAGAGTGTGGCTCTCTCTGCA 300
    |||||
Db 81 CysSerGlnIlyValValIlyValIlyAspIleSerSerGlnAspValAlaSerIleuSer 100
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAGAGGCTTCAACTCTGTTCACCTG 360
    |||||
Db 101 AspCysPheLeuThrValLeuLeuGlnIlySleuGlnIlySerPheAsnSerValHisIleu 120
QY 361 CTTGAGAGTGGGTTTGGCTGAGTCTCTGTTGTTTCTGAGCTCTGTCGAAAGAAATCC 420
    |||||
Db 121 LeuAlaGlyIlyPheAlaGlnPheSerArgCysPheProGlnIlyuGlnIlySer 140
QY 421 AGCTAGAGCTTCACTGATCTTCTGAGCTTCTGCTTACCTGCTGCAACATTTGGCCCAAC 480
    |||||
Db 141 ThrIleuValProThrCysIleSerGlnProCysIleuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTTCCAAATCTTATCTTGGTCCGACGAGATGCTTCAACAAGAGAGCTGATA 540
    |||||
Db 161 ArgIleuProAsnIleuIlyIleuGlnCysGlnAspValIleuAsnIlySleuMet 180
QY 541 CAGCAGAAATGGGATTTGTTATGTTAAATGCAAGCTTACCTGTCGAAAGCTGACTTT 600
    |||||
Db 181 GlnGlnAsnGlyIleGlyIlyValIleuAsnAlaSerAsnThrCysProIlySproAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCGTGGCTGCTGAAATGACAGCTTTTGTGAGAAATTTTG 660
    |||||
Db 201 IleProGlnIlySerHisPheLeuAlaIlyProValAlaAsnAspSerPheCysGluIlySleu 220
QY 661 CCGTGGTTGGAAGAAATGATGATTTATTTAGAGAAAGCAAGCCCTCAATGAGATGTT 720
    |||||
Db 221 ProThrPheAspIlySserValAspPheIleGlnIlyAlaIlyAlaSerAsnIlyCysVal 240
QY 721 CTAGAGCACTTTAGCTGGATCTCCGCTCGGACACATGCTATGCTTACATCATG 780
    |||||
Db 241 LeuValHisCysIleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIlyIleMet 260
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTACT 840
    |||||
Db 261 LysArgMetAspMetSerIleuAspGlnIlyArgPheValIlySgluIlySgIlyProThr 280
QY 841 ATATCTTCAACTTCAATTTCTGGGCGCAATCCCGACTATGAGAGAAATTAAGAAC 900
    |||||

```

Db 281 IleSerProAsnPheAsnIleuGlnIleuLeuAspIlyGlnIlySlyIleIlyAsn 300
 QY 901 CAGACT 906
 Db 301 GlnThr 302

RESULT 10
 ID ABB97291
 ABB97291 standard; protein; 665 AA.
 AC ABB97291;
 XX
 XX 27-JUN-2002 (first entry)
 XX
 XX

DE Novel human protein SEQ ID NO: 559.
 XX
 XX
 XX

KW Human; anti-nausea; vulnary; anti-inflammatory; immunomodulator;
 KW anti-fertility; cerebroprotective; cyostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.

OS Homo sapiens.
 PN WO20022660-A2.
 XX
 XX 21-MAR-2002.
 PD
 PF 10-SEP-2001; 2001WO-US026015.
 XX
 XX 11-SEP-2000; 2000US-00659671.
 XX
 XX (HYSE-) HYSEQ INC.
 XX

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Dermanac RT,
 XX WPI; 2002-292408/33.
 DR N-PSDB; ABB32477.
 XX

PT An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.
 XX
 XX

PS Example 2; SEQ ID NO 559; 509pp; English.

CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate haemostasis or thrombolysis e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 XX

SQ Sequence 665 AA;

Alignment Scores:

Pred. No.: 2,85e-169 Length: 665
 Score: 1552.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 1
 Query Match: 95.39% Indels: 0
 DB: 5 Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x ABB97291 (1-665)

```

QY 1 ATGGCCCATGAGATGATTTGGAAGTCAAAATGTTACTGAGAGGTGGTGGCTCTGCTGGA 60
    |||||
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
QY 61 AGTGAACGGAAGAAAGTGGCTTAATGATAGCCGGCCATTGTGGAATACAAATCATCC 120
    |||||

```

Db	21	SeerGIYhrcGIuIuYValIeulIeaSpSerArgProPheValGIuYrAenThrSer	40	XX	Key	Location/Qualifiers
QY	121	CAcATTTTGGAGCCATTATATCACTGCTCCAGGTTATGAAGGAGGTTGCAACAG	180	FT	Modified-site	12 /note= "O-phosphorylated"
Db	41	HisIleIeuGIuIaIleAenIleAenCysSerIuYleuIeUcIysArgIleuGIuIn	60	FT	Modified-site	21 /note= "O-phosphorylated"
QY	181	GACAAAGTTAAATTCAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTGAT	240	FT	Modified-site	23 /note= "O-phosphorylated"
Db	61	AspIuYValIeulIeThrcIuIeulIeGIuIeSerIaIyshiIyIuYValIeAsp	80	FT	Modified-site	38 /note= "N-glycosylated"
QY	241	TGCAGTCAGAGGTTGATTATCGATCAAAAGCTCCCAAGATGTGGCTCTCTCTCA	300	FT	Modified-site	38 /note= "N-glycosylated"
Db	81	CysSerGIuIuYValIValIValIYrAspGIuIeSerSerGIuIeValIaIeSerIeUserSer	100	FT	Modified-site	49 /note= "N-glycosylated"
QY	301	GACTGTTTTCCTGCTGCTCTCTGGGTTAACTGAGAAAGCTTCAACTCTGTTCACTG	360	FT	Modified-site	72 /note= "O-phosphorylated"
Db	101	AspCysPheIeulThrcIuIeulIeGIuIeSerArgPheProIuYleuCysGIuIYIuYSer	120	FT	Modified-site	82 /note= "O-phosphorylated"
QY	361	CTTGCAAGTGGGTTGGTGGATTCTCTCGTTGTTCCCTGGCCCTGTGAAGGAAATCC	420	FT	Region	85..298 /note= "O-phosphorylated"
Db	121	LeuAlIeGIuIYpheaIaGIuIeSerArgPheProIuYleuCysGIuIYIuYSer	140	FT	Modified-site	91 /note= "VHL-type dual specificity phosphatase signature"
QY	421	ACTCTAGTCCCTACCTGCATTTCTCAGCCCTTCTACTGTTGCCAATGGGCAACC	480	FT	Modified-site	190 /note= "O-phosphorylated"
Db	141	ThrIeUValProThrcIysIeSerGIuIeProCysIeUProValIaAenIleGIYProThr	160	FT	Modified-site	190 /note= "N-glycosylated"
QY	481	CGAATCTTCCCAATCTTTATCTTGGCTGCGACGAGATGCTCTCAACAAGAGCTGATA	540	FT	Modified-site	212 /note= "N-glycosylated"
Db	161	ArgIleIeUProAenIuYrIeUglYCYsGIuIaIyValIeUbnIySGIuIeUmer	180	FT	Modified-site	214 /note= "N-glycosylated"
QY	541	CAGCAGATGGAGTTGGTATGTGTTAAATGCCAGCATATCTGTCCAAAGCCTGACTTT	600	FT	Active-site	220..280 /note= "O-phosphorylated"
Db	181	GIuIaIeSerGIuIYIeGIYrYValIeUbnIaIeSerAenThrcIysProIuYProAenPhe	200	FT	Region	237..278 /note= "Tyrosine specific protein phosphatase"
QY	601	ATCCCGAGTCTCTATTCCTGCGTGCCTGTGATGATGACAGCTTTGTGAAATTTTG	660	FT	Modified-site	256 /note= "Y phosphatase signature"
Db	201	IleProGIuIeSerIaIePheIeUaIyValIProValIaAenPheSerPheCysGIuIYIleU	220	FT	Modified-site	280 /note= "O-phosphorylated"
QY	661	CCGTGGTGGCAAAATCAGTAAATTTTCATTGAGAAAGCAAAAGCCTTCAATGGATGTGT	720	FT	Modified-site	300 /note= "O-phosphorylated"
Db	221	ProTrIeUbnAspIySerValIaAspPheIleGIuIYValIaIySAlaSerAenGIYCYsVal	240	FT	Modified-site	369 /note= "N-glycosylated"
QY	721	CTAGTCACCTGTTAGCTGGAGTCCCGCTCCGACACATGCTATGCGCTACATCAG	780	FT	Modified-site	393 /note= "O-phosphorylated"
Db	241	LeuValIeIeCysIeUbnIaGIYIeSerArgSerIaIeThrcIeAlaIeAlaIyIleUmer	260	FT	Modified-site	421 /note= "O-phosphorylated"
QY	781	AAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT	840	FT	Modified-site	422 /note= "O-phosphorylated"
Db	261	IysArgMetAspMetSerIeUbnAspGIuIaIyArgPheValIySGIuIYsArgProThr	280	FT	Modified-site	434 /note= "O-phosphorylated"
QY	841	ATATCTTCCAAACTTCAATTTTCTGGGCGAACTCTTGAGCTATGAGAAAGATTAAAGAC	900	FT	Modified-site	439 /note= "O-phosphorylated"
Db	281	IleSerProAenPheAenPheIeUglYGIuIeUbnAspIyGIuIYsIleYsIleYsAen	300	FT	Modified-site	468 /note= "O-phosphorylated"
QY	901	CAGACT 906		FT	Modified-site	471 /note= "O-phosphorylated"
Db	301	GIuIeThr 302		FT	Modified-site	479 /note= "O-phosphorylated"
RESULT 11				FT	Modified-site	528 /note= "O-phosphorylated"
AAB20325				FT	Modified-site	590 /note= "O-phosphorylated"
ID AAB20325				FT	Modified-site	590 /note= "O-phosphorylated"
XX AAB20325;				FT	Modified-site	597 /note= "O-phosphorylated"
AC AAB20325;				FT	Modified-site	605 /note= "O-phosphorylated"
XX 29-MAY-2001 (first entry)				FT	Modified-site	610 /note= "O-phosphorylated"
DT 29-MAY-2001 (first entry)				FT	Modified-site	613 /note= "O-phosphorylated"
XX Human protein phosphatase and kinase protein-4.				FT	Modified-site	618 /note= "O-phosphorylated"
XX Protein phosphatase and kinase protein, pPKR-4; human;				FT	Modified-site	628 /note= "O-phosphorylated"
KW gastrocnemius disorder; immune system disorder; neurological disorder;				FT	Modified-site	641 /note= "O-phosphorylated"
KW cell proliferative disorder; cancer; diagnosis; therapy.				FT	Modified-site	
XX Homo sapiens.				FT	Modified-site	
OS				FT	Modified-site	

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;
XX WPI; 2003-395539/38.
DR N-PSDB; ADA53105.
XX
PT New polynucleotide encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
XX Claim 14; SEQ ID NO 2312; 205bp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA5433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 665 AA;

Alignment Scores:
Pred. No.: 1,83e-168 Length: 665
Score: 1545.00 Matches: 299
Percent Similarity: 99.34% Conservative: 1
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 94.96% Indels: 0
DB: Gaps: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x ADA54744 (1-665)

QY 1 ATGCGCCATGAGAGTGTGGAATCAATTTGTACTGAGAGGTGGCTGCTGCGAA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGln 20
QY 61 AGTGAACGGAAGAAAGTGTGCTGATTTGATGAGCGGCGCAATTTGTGAATACATACATCC 120
DB 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnIleSer 40
QY 121 CACATTTTGGAAAGCATTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 180
DB 41 HisIleLeuGlnAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleGln 60
QY 181 GACAAAGTGTATTTACAGAGCTCATCCAGCATTCACGCAACATTAAGTTGACATTAAT 240
DB 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATGTTAGCATCAAGCTCCAGAGTGTGCTGCTCTCTTCA 300
DB 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTCACTGACTTCTGGGTAAACCTGGAAGAGCTTCACTCTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGGTGGGTGGTGGTCTGCTCTGTTTCCCTGGGCTCTGTGAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
QY 421 ACTCTGATCCCTACCGCATTTCTCAGCGCTGCTTACCTGGTGGCAAACTGGGCGCAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGACGAGATGCTCTCAACAAGAGCTGATA 540
DB 161 LeuIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuMet 180
QY 541 CAGCAAGAGGAGTGGTATGTTATGTTAAATGCAAGTATACCTGTCCAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG 660

DB 201 IleProGlnSerHisPheLeuArgValProValAlaAspSerPheCysGluLysIleLeu 220
QY 661 CCGTGGTGGACAAATCAGTATGATTTTCATTGAGAAAGCAAAAGCTTCAATGATGTCTT 720
DB 221 ProTyrLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTATGCACTGTTTGGTGGATCTCCGCTCCGACCATGCTATGCTGCTTACATCAAG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAACACTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGluLysArgProThr 280
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTCGCATTAAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlnLeuLeuAspTyrGluLysLysIleLysAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302

RESULT 13
AAU79161
ID AAU79161 standard; protein; 665 AA.
XX
AC AAU79161;
XX
XX 02-JUL-2002 (first entry)
DE Human dual-specificity phosphatase-3 (DSP-16) mutant protein #1.
XX
XX Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
XX mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
XX cancer; graft-versus-host disease; allergy; metabolic disease;
XX abnormal cell growth; abnormal cell proliferation; contact inhibition;
XX cell cycle abnormality; anchorage independent cell growth; apoptosis;
XX intercellular adhesion; DSP-16 modulator; mutant; mutain.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
FH Key location/Qualifiers
FT Misc-difference 213 /note= "wild-type Asp substituted by Ala"
XX
XX PN WO200226997-A2.
XX
XX PD 04-APR-2002.
XX
XX PF 25-SEP-2001; 2001WO-US030124.
XX
XX PR 26-SEP-2000; 2000US-0235487P.
XX
XX PA (CEPT-) CEPTYR INC.
XX
XX PI Luche RM, Wei B;
XX
XX DR WPI; 2002-315802/35.
XX
PT New DSP-16 polypeptide, useful for identifying modulators of its
PT activity, which can be used in the treatment of disorders such as
PT Duchenne muscular dystrophy, or cancer.
XX
XX PS Claim 46; Page; 87pp; English.
XX
XX The present invention relates to a new polypeptide, DSP-16, having a 665
XX amino acid sequence, given in the specification, or a variant having at
XX least 50 % identical residues, which retains the ability to
XX dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
XX invention can be used for identifying agents which modulate DSP-16
XX activity, for modulation of a proliferative response in a cell, survival

CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne
CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
CC cell proliferation, and cell cycle abnormalities. The present anti-no acid
CC mutant protein #1. Note: This sequence is not shown in the specification
CC but is derived from the wild-type human DSP-16 (AAU79156) protein given
CC in figure 2 of the specification
XX
SQ Sequence 665 AA;

Alignment Scores:

Pred. No.:	2,38e-168	Length:	665
Score:	1544.00	Matches:	299
Percent Similarity:	99.34%	Conservative:	1
Best Local Similarity:	99.01%	Mismatches:	2
Query Match:	94.90%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) x AAU79161 (1-665)

```
QY 1 ATGGCCCATGAGATGATGGAAGTCAATGTTCTGAGAGGTTGGCTGCTGCGAA 60
Db 1 MetAlaHLeuMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
QY 61 AGTGAACGGAAGAAAGTCTGCTAATGATACCGCGCATTTGTGAAATCAATACATCC 120
Db 21 SerGlyThrGluValValLeuLeuIleAspSerArgProPheValGluThrSerThrSer 40
QY 121 CACATTTTGAAGCATTATATCACTGCTCCAGCTTATGAGAGGAGGTTGCAACG 180
Db 41 HisIleuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgTrgLeuGlnGln 60
QY 181 GACAAAGTGTATTAATACAGAGCTCAGCATTCAGCAAAACATTAAGTTGACATTGAT 240
Db 61 AspLysValLeuIleThrGlnIleGlnHisSerAlaLysHisLysValAspIleLeu 80
QY 241 TGCAAGTCAAGAGTTGATGTTTACGATCAAGAGCTCCAGAGTTGCTCTCTCTTCA 300
Db 81 CysSerLysValValValValValValValValValValValValValValValValVal 100
QY 301 GACTGTTTCTCACTGATCTTGGGTAACCTGAGAGAGAGGTTCACTCTGTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTGGCAGGCGGTTTCTGAGTTCTCTGTTGTTCCCTGGCTCTGAGAGAGAGAAATCC 420
Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLysGluGlyLysSer 140
QY 421 ACTAGTCCCTACCGCATTTCTCAGCCTTCTTACCTGTTGCAACATTTGGCCACAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTCCAGAGAGATGCTCCCAAGAGAGAGATGA 540
Db 161 ArgIleLeuProAsnLeuLysLeuGlyCysGlnArgSerValLeuAsnLysGlnLeuMet 180
QY 541 CAGCAGAAATGGGATTTGTTATGTTTAAATGCAAGCTATCTGTTCAAAAGCTTGAATT 600
Db 181 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGGCTCATTTCTGCGCTGCGCTGATGATGACAGCTTTTGGAGAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAlaSerPheCysGlnLysIleLeu 220
QY 661 CCGTGTGTGACAAATAGTAGATTTCATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 221 ProThrPheAspLysSerValAspPheIleGlyLysValAlaLysValAsnGlyCysVal 240
QY 721 CTAGTGACGCTTTAGCTGGAGATCTCCGCTCCGACCATCGCATGCTATCATCATG 780
```

```
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGATGAGACATGCTTTTATGATGACGATTTAGATTTTGAAGAGAGAGAGAGAGAG 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCCTGGCATATGAGAGAGAGATTAAAGAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspLysIleLysIleLysAsn 300
QY 901 CAGACT 906
Db 301 GlnThr 302
RESULT 14
ID ABR52352 standard; protein; 665 AA.
AC ABR52352;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein relating to the invention SEQ ID NO: 42.
XX
XX antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antiproliferative; cardiac; cytosolic; gene therapy; liver disease;
XX proliferative disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX Homo sapiens.
XX
XX WO200257460-A2.
XX
XX 25-JUL-2002.
XX
XX 20-DEC-2001; 2001WO-US050459.
XX
XX 20-DEC-2000; 2000US-025686P.
XX 30-MAR-2001; 2001US-028018P.
XX 01-MAY-2001; 2001US-0287735P.
XX 05-JUN-2001; 2001US-0295848P.
XX 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Jackson DG, Feder J, Nelson T, Mintier G, Ramamathan C, Lee L,
XX Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
XX Krystek S, Mcatee P, Suchard S, Banas D;
XX WPI; 2002-599721/64.
XX N-PSDB; ACC60521.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
XX the prevention or treatment of e.g. proliferative and cardiovascular
XX disorders.
XX
XX Claim 5; Fig 5; 801pp; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
XX polynucleotide having a nucleotide sequence selected from 40
XX polynucleotides fully defined in the specification. The polynucleotide of
XX antiarthritic, antiproliferative, hepatotropic, nephrotropic,
XX antiproliferative, cardiac, and cytosolic activity. The
XX polynucleotide may have a use in gene therapy. A polynucleotide or
XX ameliorating a medical condition, e.g. a proliferative disorder. They are
XX also useful for treating e.g. liver disease, renal failure, immunological
XX disorders including arthritis and psoriasis, cardiovascular disorders
XX such as congenital heart defects and congestive heart failure, and
XX cancer. A method of the invention is useful for diagnosing a pathological
XX condition or susceptibility to a condition in a subject. The present
```


CC sequence is used in the exemplification of the invention
XX
SQ Sequence 665 AA:

Alignment Scores:

Pred. No.:	Length:	Matches:
Score:	2,386-168	665
Percent Similarity:	1544.00	299
Best Local Similarity:	99.34%	1
Query Match:	99.01%	2
DB:	94.90%	0
	Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) x ABR52352 (1-665)

```

QY 1 ATGGCCCAATGAGTATGGAAGTCAATGTTACTGAGAGGTTGGTGGCTTCTGCTGAA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTGTCTAATGATGAGCGGCGCATTTGTGGAATACATACATCC 120
DB 21 SerGlyThrGluValValLeuLeuIleAspSerArgProPheValGlnThrSer 40
QY 121 CACATTTGGAGACCATTAATATCACTGCTCCAGGTTATGAAGCGAAGTTGCAACAG 180
DB 41 HisIleLeuGlnIleAlaIleAsnIleAsnGlySerLeuMetLeuArgArgLeuGln 60
QY 181 GACAAAGTGTATTTACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
DB 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAlaPheIleAsp 80
QY 241 TGGAGCCAGAGGTTGATGATTACGATCAAGCTCCCAAGATGTTGGCTCTCTCTTCA 300
DB 81 CysSerGlnLysValValValIleValIleValIleValIleValIleValIleValIleVal 100
QY 301 GACTGTTTCTCAGTCTACTTCTGGGTAAACTGAGAGAGAGCTTCACTCTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisIleLeu 120
QY 361 CTGGCAGGTTGGTTTGGTGGTCTCTGCTGTTTCCCTGGCTCTGTGAGAGAAATGC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProLysLeuGlyLysSer 140
QY 421 ACTCTAGTCCCTTCTGATCTTCTGACGCTTGTCTTACCTGTTGCCAATTTGGGCAACC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTCTTCCCAATTTTATCTTGGCTGCGCAGCAGAGATGCTCTCAACAAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuThrLeuGlyCysGlnArgAspValLeuAsnLysGlnLeuMet 180
QY 541 CAGCAGATGGGATTTGGTTATGTTTAATGCCAGCTATTCCTGTCGAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyThrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCCGGGTGGGCTCGTAATGACAGCCTTTGTGAGAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
QY 661 CCGTGGTGGACAAATCAGTAGATTTTCATTGAGAGAAAGCAAAAGCTTCAATGATGATGTT 720
DB 221 ProThrLeuAspLysSerValAlaPheIleGlnLysValAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTAGTGCATCTTTAGCTGGGATCTCCCGTCCGCCCACTTCGCTTATCGCTTATCATCAG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaLysIleMet 260
QY 781 AAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAAGAAAAAACCTTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaLysArgPheValIleGlnLysArgProThr 280
QY 841 ATATCTCAAACTTCAATTTCTGGGCGCACTCTGAGCATATGAGAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAlaLysGlnLysLysIleLysAsn 300

```

```

QY 901 CAGACT 906
DB 301 GlnThr 302

RESULT 15
ID AAM25744
AAM25744 standard; protein; 672 AA.
AC AAM25744;
XX 16-OCT-2001 (first entry)
DT XX
DE Human protein sequence SEQ ID NO:1259.

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW anti-inflammatory; antirheumatic; antidiabetic; immunosuppressive;
KW antibacterial; endocrine; cardiac; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiparasitic; haemostatic; vulnery; antidiabetic; osteopathic; eczema;
KW dermatological; antiallergic; antistatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.

XX Homo sapiens.
OS
XX WO200153455-A2.
PN
XX 26-JUL-2001.
PD
XX 22-DEC-2000; 2000WO-US035017.
PF
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
XX
XX (HYSR-) HYSRQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI WPI; 2001-457603/49.
XX DR N-PSDB; AAM99685.
XX
XX Isolated human polynucleotides encoding polypeptides, useful for the
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
XX Claim 20; Page 260; 1217pp; English.
XX
XX AAM99166 to AAM99904 encode the human proteins given in AAM25725 to
XX AAM25963. The proteins can have activities based on the tissues and cells
XX they are expressed in, such as: antineoplastic; antirheumatic;
XX antidiabetic; immunosuppressive; antibacterial; endocrine; cardiac;
XX central nervous system; virucide; anti-HIV; fungicide; antimutagen;
XX cardiovascular; antianaemic; antiparasitic; haemostatic; vulnery;
XX antidiabetic; osteopathic; dermatological; antiallergic; antistatic;
XX antiparkinsonian; cytostatic; neuroprotective; antidepressant; nootropic;
XX encoding them can be used in gene therapy, antisense therapy and vaccine
XX production. The proteins and polynucleotides are useful for screening for
XX agonists or antagonists of a protein and for the treatment and diagnosis
XX of disorders associated with the activity of a protein e.g. inflammation,
XX rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
XX neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
XX infections, autoimmunity, genetic diseases, haematopoietic disorders,
XX anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
XX osteoporosis, severe combined immunodeficiency, eczema, allergic

```

CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX
SQ Sequence 672 AA;

Alignment Scores:

Pred. No.:	2,39e-168	Length:	672
Score:	1544.00	Matches:	299
Percent Similarity:	99.34%	Conservative:	1
Best Local Similarity:	99.01%	Mismatches:	2
Query Match:	94.90%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) x AAM25744 (1-672)

QY 1 ATGGCCCATGATGATGATGAACTCAAAATTGTTACTGAGAGTTGGTCTGCTGAA 60
Db 8 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgGlyValAlaLeuLeuGlu 27
QY 61 AGTGAACCGAAGAAAGTGTCTGCTAAATTGATGCCGCCATTGTGGAATCAATACATCC 120
Db 28 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 47
QY 121 CACATTTTGAAGACCATTAATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 180
Db 48 HisIleuGluValIleAlaHisIleAsnGlySerIleuMetLysArgLeuGlnGln 67
QY 181 GACAAAGTGAATTAATTAACAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
Db 68 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaHisIleValAspIleAsp 87
QY 241 TGCAGTCAGAGTTGATGATTAAGATCAAGCTCCCAAGTTGCTCTCTCTTCA 300
Db 88 CysSerGlnLysValValIleValIleValIleValIleValIleValIleValIleVal 107
QY 301 GACTGTTTCTCACTGACTCTTGGGTAACCTGAGAGAGCTTCAACTGTTGACCTG 360
Db 108 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisIleu 127
QY 361 CTTGCAAGTGGGTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTAAGAAATCC 420
Db 128 LeuAlaGlyIlePheAlaGlnPheSerArgCysPheProGlyLeuGlyGlySer 147
QY 421 ACTCTAGTCCCTCACTGCTGCTTCTGAGCTTCTTACCTGTTGCCAATGAGGCAACC 480
Db 148 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 167
QY 481 CGAATTCCTCCCAATCTTATCTTGGTCCAGCGAGATGTCCTCAACAAAGAGCTGATA 540
Db 168 ArgIleLeuProAsnLeuIleuGlyCysGlnArgAspValLeuAsnLysGluLeuMet 187
QY 541 CAGCGAATGGGATGGTTATGTGTTAAATGCCAGCTATACCTGTCCAAGCTGACTTT 600
Db 188 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerAsnThrCysProLysProAspPhe 207
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGTAATGAGAGCTTGTGTGAAAATTTTG 660
Db 208 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 227
QY 661 CCGTGGTTGACAAATCATGATGATTTATTTAGAGAAAGCAAAAGCTCCAAATGATGTT 720
Db 228 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnLysCysVal 247
QY 721 CTAGTGACTGTGTAGCTGGGATCTCCCGCTCGGCACCATGCTATGCTTACATCATG 780
Db 248 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 267
QY 781 AAGAGATGACATGCTCTTATGATGAAGCTTACAGATTGTGAAAAGAAAAGCTTACT 840
Db 268 LysArgMetCysPheCysLeuAspGluAlaIleArgPheValLysGluLysArgProThr 287
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGGACTATGAGAGAGATTAAAGAC 900

Db	288	IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysLysIleLysAsn 307
QY	901	CAGACT 906
Db	308	GlnThr 309

Search completed: June 21, 2004, 12:34:22
Job time : 51.4105 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:40:31 ; Search time 43.0982 Seconds
(without alignments)
11869.448 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443
Perfect score: 1627
1 atggccacgagatgatg9.....agaagattaagaaccagact 906

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Egapop 6.0 , Egapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1163542 seqs, 282313646 residues
Total number of hits satisfying chosen parameters: 2327084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=xip
-Q/cgn2_1/USPTO.spool.p/US10029345/runat.21062004.122818.4331/app.query.fasta_1.3278
-DB=Published Applications AA -OPT=fasta -SUPR=trabp -MINMATCH=0.1
-LOPCU=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR.SCORE=pcr -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10029345 @CGN 1 21 @runat.21062004.122818.4331
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	1552	95.4	662	12	US-10-072-012-258	Sequence 258, App
2	1552	95.4	665	9	US-09-816-494-2	Sequence 2, Appl1
3	1552	95.4	665	9	US-09-964-277-2	Sequence 2, Appl1
4	1552	95.4	665	12	US-10-072-012-680	Sequence 680, App
5	1552	95.4	665	12	US-10-168-506-14	Sequence 14, Appl
6	1552	95.4	665	12	US-10-343-357-7	Sequence 7, Appl1
7	1552	95.4	665	15	US-10-377-072-26	Sequence 26, Appl
8	1552	95.4	665	16	US-10-257-026-2	Sequence 2, Appl1
9	1552	95.4	665	16	US-10-648-593-240	Sequence 240, App
10	1552	95.4	665	16	US-10-648-593-247	Sequence 247, App
11	1552	95.4	665	16	US-10-648-593-247	Sequence 679, App
12	1552	95.4	690	12	US-10-072-012-679	Sequence 5203, A
13	1552	95.4	690	12	US-10-072-012-703	Sequence 703, App
14	1552	95.4	665	12	US-10-425-012-54204	Sequence 681, App
15	1545	95.0	665	12	US-10-072-012-681	Sequence 2312, Ap
16	1545	95.0	665	15	US-10-094-749-2312	Sequence 1259, Ap
17	1517	93.2	672	12	US-10-296-115-1259	Sequence 682, App
18	1514	93.1	677	12	US-10-072-012-682	Sequence 683, App
19	1504.5	92.5	680	12	US-10-072-012-256	Sequence 256, App
20	988	60.7	625	12	US-10-072-012-699	Sequence 699, App
21	985	60.5	663	12	US-10-072-012-700	Sequence 700, App
22	837	51.4	616	12	US-10-072-012-266	Sequence 266, App
23	807	49.6	155	9	US-09-964-277-7	Sequence 7, Appl1
24	740	45.5	517	9	US-09-964-277-21	Sequence 21, Appl
25	654.5	40.2	169	14	US-10-346-356-15	Sequence 15, Appl
26	654.5	40.2	170	9	US-09-775-925-26	Sequence 26, Appl
27	654.5	40.2	170	9	US-09-847-519A-11	Sequence 16, Appl
28	654.5	40.2	170	12	US-10-655-073-16	Sequence 16, Appl
29	654.5	40.2	170	14	US-10-314-058-14	Sequence 14, Appl
30	654.5	40.2	170	14	US-10-405-808-16	Sequence 16, Appl
31	631.5	38.1	461	12	US-10-072-012-701	Sequence 701, App
32	620	38.1	501	12	US-10-072-012-702	Sequence 702, App
33	619	38.0	155	9	US-09-964-277-6	Sequence 6, Appl1
34	619	38.0	155	9	US-09-955-732-6	Sequence 6, Appl1
35	469	28.8	444	9	US-09-964-899-47	Sequence 47, Appl
36	469	28.8	482	12	US-10-058-270A-130	Sequence 130, App
37	469	28.8	482	12	US-10-346-356-2	Sequence 2, Appl1
38	469	28.8	482	16	US-10-648-593-164	Sequence 164, App
39	462.5	28.4	381	14	US-10-184-832-2	Sequence 2, Appl1
40	428	26.3	334	12	US-10-060-065-25	Sequence 25, Appl
41	428	26.3	334	14	US-10-059-585-46	Sequence 46, Appl
42	427	26.2	334	14	US-09-736-457-805	Sequence 805, App
43	427	26.2	334	9	US-09-736-457-827	Sequence 827, App
44	427	26.2	394	9	US-09-902-941-805	Sequence 805, App
45	427	26.2	394	9	US-09-902-941-827	Sequence 827, App

ALIGNMENTS

RESULT 1
US-10-072-012-258
Sequence 258, Application US/10072012
Publication No. US2004003493A1
GENERAL INFORMATION:
APPLICANT: Tcheinev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zehnusen, Bryan
APPLICANT: Paturajan, Weera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Babha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Futak, Katarzyna
APPLICANT: Grose, William M.

APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 258
LENGTH: 662
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-258

Alignment Scores:

Pred. No.: 6,02e-152 Length: 662
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-072-012-258 (1-662)

QY 1 ATGGCCATGAGATGATGGAATGGAATGTTTACTGAGAGGTTGGCTGCTGGAA 60
DB 1 MetAlahisglumetilegilythrghinilevalthrgluargluvalAlaleuLeuGlu 20
QY 61 AGTGAACGGAAGAGTGTGCTATTTGATAGCGGCGCATTTGTGAATCAATACATCC 120
DB 21 SerGlyThrgluValleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
QY 121 CACATTTTGAAGCATTATATATCAACTGCTCCAAAGCTTGTGAAGGAAGTTGCAACG 180
DB 41 HisileuLeuGluAlaleasnileasnileasnileasnileasnileasnileasnile 60
QY 181 GACAAAGTGTAAATTAAGAGCTCATCCAGATTCAAGCAACATTAAGTTGACATTGAT 240
DB 61 AspIysValleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
QY 241 TGCGTCGAGAGGTTGATGTTTACATCAAGCTCCCAAGATGTTGCTCTCTTCA 300
DB 81 CysSerGlnIysValValValIyrAspGlnSerSerGlnAspValAlaserIeuSer 100
QY 301 GACTGTTTCTCACTGATCTTGGGGTAACTGAGAGAGGTTCACTGTTCACTG 360
DB 101 AspCysPheLeuThrValleuLeuGlyIysLeuGluIysSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGGTGGTTTCTAGATTCTCTCTGTTTCCCTGGGCTCTGGAAGAAATTC 420

DB 121 LeuAlaGlyIyIyPheAlaGluPheSerArgCysPheProGlyLeuCyGluGlyIyLysSer 140
QY 421 ACTTACTGCTTACCTGATGATTTCTGAGCTTGTCTTACCTGTTGGCAACTTGGCCCAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuPheProValAlaAsnIleGlyProThr 160
QY 481 CGAATTCCTCCCAATCTTATTTATTTGGCTGCGCAGAGAGTGTCTCAACAAGAGCTGAT 600
DB 161 ArgIleLeuPheAsnLeuIyruendlyCysGlnIyrAspValleuAsnIySgIuLeuMet 180
QY 541 CAGCGAATGGGATGTTATGTGTAAATGCCGATATACCTGTCCAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyIyValleuAsnAlaserAsnThrCysProIysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGGTCGCTGATGATGACGCTTTGTGAGAAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIySleu 220
QY 661 CCGGTGTTGACCAATCATGATGATTTCAATTGAGAAAGCAAGCCCAATGATGTGT 720
DB 221 ProIleuAspIySserValAspPheIleGlyIySalIyAlaserAsnIyCysVal 240
QY 721 CTAGTCACTGTTTACGCTGAGATCTCCGCTCCGACCATGCTATGCTTATGCTTACATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerIyrSerAlaThrIleAlaIleAlaIyIleMet 260
QY 781 AAGAGATGAGACATGCTTTTATGATGAGCTTACAGATTTGTGAAGAAAAGCACTACT 840
DB 261 LysArgMetAspMetSerIeuAspGluAlaIyIyrArgPheValIySgIuIyAspProThr 280
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCCCTGACATGATGAGAAAGATTAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyIyIeuLeuAspIyrgluIySlyIleIyAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302

RESULT 2

US-09-816-494-2
Sequence 2, Application US/09816494
Patent No. US20020034807A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-09-816-494-2

Alignment Scores:

Pred. No.: 6,03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-09-816-494-2 (1-665)

QY 1 ATGGCCATGAGATGATGGAATGGAATGTTTACTGAGAGGTTGGCTGCTGCGAA 60
DB 1 MetAlahisglumetilegilythrghinilevalthrgluargluvalAlaleuLeuGlu 20

```

OY      61  AGTGAACGAAAAAGTGCCTGAATTAATGATAGCGGCACATTTGTGAATAACAATACATCC 120
Db      21  SerGlythnGlnLysValLeuLeuIleAepSerArgProPheValGlnLysrhnntnSer 40
OY      121  CACATTTTGGAAAGCCATTAAATATCAACTCTCCCAACCTTATGAAAGCAAGGTTCACAAG 180
Db      41  HstIleLeuGlnLalaleanIleahenCyserLysLeuMetLysrhrgrLeuGlnGln 60
OY      181  GACAAAGTGTAAATPACAGAGCTCATCCAGCATTCAGCGAAACATPAAGGTTCAGATTGAT 240
Db      61  AapLysValLeuIlethrGlnLeuIleGlnHisSerAlaLysHisLysValAapIleAap 80
OY      241  TGCACATCGAAGGTTGACTTTATACGATCAAAAGTCCCAAGATCTCCCTCTCTTTCA 300
Db      81  CysSerGlnLysValValValTyrrAepGlnSerSerGlnAapValAlaSerLeuSerSer 100
OY      301  GACTGTTTTCCTCAGCTGTACTCTGGGGTAAACTGGAAGAAGCTTCAACTCTGTTCACCTG 360
Db      101  AapCySpheLeuThrValLeuLeuGlnLysLysLeuGlnLysSerPheAmservaIhLsLeu 120
OY      361  CTTGCAGAGTGGGTTTGCTGAGTTCTCTCGTTGTTTCCTCGCCCTCTGTGAAGAAATCC 420
Db      121  LeuAlaGlyGlyPheAlaGlnPheSerArgCySpheProGlyLeuCySglnGlyLysSer 140
OY      421  ACTCTAGTCCCTCACTCGCATTTTCTCAGCCTTGTTACTCTGTTCGCAACATTTGGCCAAAC 480
Db      141  ThrLeuValProThrCySrlleSerGlnProCySleuProValAlaAasnIleGlyProThr 160
OY      481  CGAATTCCTCCCAATCTTATCTTGAGCTGACGCGAGATGTCTCAACAAGAGAGCTGATA 540
Db      161  ArgIleLeuProAmsLeuTyrrLeuGlyCySglnArgAapValLeuAmsLysGlnLeuMet 180
OY      541  CAGCAGAAATGGGATTGGTTATGTGTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 600
Db      181  GlnGlnAmsGlyIleGlyTyrrValLeuAmsAlaSerAasnThrCySpProLysProAapPhe 200
OY      601  ATCCCGAGTCTCATTTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db      201  IleProGlnSerhIshPheLeuArgValProValAasnAapSerPheCySglnLysIleLeu 220
OY      661  CGGTGGTTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCCAATGAGATGTTGTT 720
Db      221  ProThrLeuAapLysSerValAapPheIleGlnLysAlaLysAlaSerhngLysVal 240
OY      721  CTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGACCATGCGTTATGCGCTACATCATG 780
Db      241  LeuValhIshCyLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrrIleMet 260
OY      781  AAGAGGATGAGACATGCTTTTAAATGAGCTTACAGATTGTGTGAAGAAAAAGACCTTACT 840
Db      261  LysArgMetAapMetSerLeuAapGlnAlaTyrrArgPheValLysGlnLysArgProThr 280
OY      841  ATATCTCCAAACTTCATTTTCTGGGCCCAACTCTTGACATATGAGAAAGATTAAGAAC 900
Db      281  IleSerProAapPheAasnPheLeuGlnLeuLeuAapTyrrIleLysLysIleLysAasn 300
OY      901  CAGACT 906
Db      301  GlnThr 302

RESULT 3
US-09-964-277-2
; Sequence 2, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125_434
; CURRENT APPLICATION NUMBER: US/09/964_277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq For Windows Version 4.0

```

```

1 | SEQ ID NO 2
2 | LENGTH: 665
3 | TYPE: PRT
4 | ORGANISM: Homo sapiens
5 | US-09-964-277-2

Alignment Scores:
Pred. No.: 6,03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-964-277-2 (1-665)

QY 1 ATGGCCCATGAGATGATGGAAGCTCAAAATTGTTACTGAGAGGTTGGTGGCTCTGCTGAA 60
   |||||
Db 1 Metalaahlsblumellieglythrghnllevalthngluargleuvalalaleuengl 20

QY 61 AGTGAACGGAAAAAGTGTGCTGTAATTGATAGCCGGCCATTGTGGAATACATATACATCC 120
   |||||
Db 21 SerglythngluylvalleuvalleuuleileapseraragprophvalglurAenthSer 40

QY 121 CACATTTTGGAAAGCCATTAAATCAACTGCTCCAGCTTATGAAGCGAAGCTTGCAACG 180
   |||||
Db 41 HlsileuenglualaleasnilleancyserylslleuMetlysrarglrleugln 60

QY 181 GACAAAGTGTAAATACAGAGCTCACTCCGCAATTCAGCGAAACATPAAGGTGACATTGAT 240
   |||||
Db 61 AsplyvalleuullethngluuuleeglnhlsesrAlalyhlsvalaplleap 80

QY 241 TGCAGTCAGAAGGTGTGATGTTTACGATCAAAAGCTCCCAAGATGTTGCCCTCTCTCTCA 300
   |||||
Db 81 CyssergluylvalvalValValYrnapglinserserglnaspvalAlaserleuSer 100

QY 301 GACTGTTTCTACACTGACTTCTGGGTAACCTGAGAGAAGACTTCAACTCTCTTCACTG 360
   |||||
Db 101 AspCysphelenthValleuengllylsleugluylsSerpheasnserValhlsleu 120

QY 361 CTGGAAGGAGGTTGCTGAGTTCCTCGTGTTCCTCGTGTTCCTCGGAGCTGTGAAGAAATCC 420
   |||||
Db 121 LeuAlaGlyGlyPheAlaGluPheSerargCyspheproGlyIleuCysegluGlylsSer 140

QY 421 ACTTAGTCCCTCACTGCACTTCTCAGACCTTGCCTTACCTGTTGCCAACAATGGGCCAAC 480
   |||||
Db 141 ThrleuValProthrCyselIeserGlnProCysleuProvalAlaasnilleGlyProthr 160

QY 481 CGAATTCTTCCCAATCTTTATTTGGCTGCGCAGCGAGATGTCTCAACAAGAGCTGAT 540
   |||||
Db 161 ArgIleleuProAsnleuIyrlleuGlyCysGlnAgsaspValleuasnlysgluuMet 180

QY 541 CAGCAGATGGGATGTTAGTGTAAATGCGAGTAACTCGTCCAAAGCTGACTT 600
   |||||
Db 181 GlnGlnaenglylleglyrValleuasnAlaseranthCysProlyspRoasphe 200

QY 601 ATCCCGAGTCTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTAGAAAAATTTTG 660
   |||||
Db 201 IlleproGluSerHisphelenuArgValProvalaAsnSerPheCysgluylsIleu 220

QY 661 CCGTGGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCATGGATGTGTT 720
   |||||
Db 221 ProTrpleuAsplysSerValAspPheIllegluylsAlaIyAlaseranGlyCysVal 240

QY 721 CTAGTGCAGCTGTTTAGCTGGGATCTCCGCGTCCGGCCAGCATGGCTATGCCCTACATCAG 780
   |||||
Db 241 LeuValHisCysleuAlleglylleserargserAlaathlIsAlaIleAlaIyrlIleMet 260

QY 781 AAGAGAGATGACATGCTCTTGTAGATCAAGCTTACAGATTTGTGAAAGAAAAAGACCTACT 840
   |||||
Db 261 LysaArgMetAspMetSerleuAspGluAlaYrArgPheValIyGluylsArgProthr 280

QY 841 ATATCTCCAAACTCAATTTTCTGGGCCCACTCTGGACTATAGAGAAGAAATTAAAGAC 900

```

Db 281 IleserProasnpheleuGlInleuLeuAspTyGILuLyIsleuysasn 300
QY 901 CAGACT 906
Db 301 GlnThr 302

RESULT 4

US-10-072-012-680
; Sequence 680, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:

; APPLICANT: Tcherev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Futak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 680
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-680

Alignment Scores:
Pred. No.: 6,03e-152
Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%

Length: 665
Matches: 300
Conservative: 1
Mismatches: 1

Query Match:

DB: 95.39%
12

Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-072-012-680 (1-665)

QY 1 ATGGCCCATGATGATTTGGAATCTCAAAATTGTTACTGAGAGGTTGTCGCTGCGGAA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGlu 20
QY 61 AGTGAACCGGAAAAAGTCGTGCTAATTGATGACCGGCGCATTTGGCAATTCATATCATCC 120
Db 21 SerGlyThrGlnValValLeuLeuIleAspSerArgProPheValGlnTyrsThrSer 40
QY 121 CACATTTTGAAGGCATTAATATCACTGCTCCAGAGCTTAATGAAGCAAGTTGCAACG 180
Db 41 HisIleuGlnAlaIleAsnIleAsnCySerIleuMetIysArgIleuGln 60
QY 181 GACAAAGCTTAATTACAGAGCTCATTCAGCATTCAGCAACATTAAGTTGACATTGAT 240
Db 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysValAspIleAsp 80
QY 241 TGCAGTCAGAAAGTTGATGTTACGATCAAAAGCTCCAAAGATTTGCTCTCTCTTCA 300
Db 81 CysSerGlnLysValValValValTyraSerGlnSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTCACTGCTACTTCTGCGTAACTCGAGAAAGCACTTCACTGTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGGTGGGTGCTGAGTTCTCTGTTGTTTCCGCTGCTCTGTGAAGAAATCC 420
Db 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlySer 140
QY 421 ACTTAGTCCCTCACTGATTTCTCAGCTTCTGCTTACTGTTGCCAAATGGGCCAACC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTCCTCCCAATCTTTATCTTGAGTGCACGAGATGTCCTCAACAAAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuTyrlleuGlyCysGlnArgAspValLeuAsnLysGluMet 180
QY 541 CAGCAGAAATGGGATGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 600
Db 181 GlnIleAsnGlyIleGlyTyrlleuValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTAATTTCTGCGTGTGCTGCTGATGACAGCTTTTGGAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
QY 661 CCGTGTGGCAAAATCAGTATGATTTGATGAAAGCAAAAGCTTCGATGATGTT 720
Db 221 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyVal 240
QY 721 CTAGTCACCTGTTTATGCTGGGATCTCCGCTCCGCAACATTCGCTTACATGATGATG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260
QY 781 AAGAGAGTGAACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaTyraArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAAATCTCATTTTCTGGGCAACTCTTGATGATGAGAAAGATTAAGAC 900
Db 281 IleserProasnpheleuGlInleuLeuAspTyGILuLyIsleuysasn 300
QY 901 CAGACT 906
Db 301 GlnThr 302

RESULT 5
US-10-168-506-14
; Sequence 14, Application US/10168506
; Publication No. US20040053229A1

```
/ GENERAL INFORMATION:
/ APPLICANT: PLOMMAN, GREGORY D.
/ APPLICANT: MARTINEZ, RICARDO
/ APPLICANT: WHITE, DAVID
/ APPLICANT: MANNING, GERARD
/ APPLICANT: SUDARSANAM, SUCHA
/ APPLICANT: HILL, RON
/ APPLICANT: PLANAGAN, PETER
/ TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
/ FILE REFERENCE: 038602/1351
/ CURRENT FILING DATE: 2002-06-21
/ PRIOR APPLICATION NUMBER: PCT/US00/34736
/ PRIOR FILING DATE: 2000-12-21
/ NUMBER OF SEQ ID NOS: 76
/ SOFTWARE: Patent Ver. 2.1
/ SEQ ID NO 14
/ LENGTH: 665
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-168-506-14

Alignment Scores:
Pred. No.: 6, 03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 0
Query Match: 95.39% Indels: 0
DB: 12 Gaps: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x US-10-168-506-14 (1-665)

QY 1 ATGGCCCATGAGTGAATGGAATCAATGTTACTGAGAGGTTGGTCTGCTGGA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAAGGAAAGAGTGTGCTTAATGATGAGCGGCAATTTGTGAATCAATACATC 120
DB 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
QY 121 CACATTTGGAGCCATTAATATCACTGCTCCAGCTTAATGAAGGAAAGTTGCAAC 180
DB 41 HisIleLeuGluValIleLeuIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
QY 181 GACAAAGTGTAAATCAAGAGCTACCCAGCATTCAGCGAAATAAGTTGACATTGAT 240
DB 61 AspLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
QY 241 TGCAGTCAGAGGTTGATGATGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
DB 81 CysSerGlnLysValValValValValValValValValValValValValValVal 100
QY 301 GACTGTTTCTCACTGACTTCTGGGTAAACTGGAGAAGAGCTTCACTCTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheLeuValHisLeu 120
QY 361 CTTCGAGGAGGTTTGGTGAATCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlySer 140
QY 421 ACTGTAGCCCTACCTGCAATTTCTGAGCTGTGCTTACTGTTCCAAACTGGGCAACC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTCCTCCCAATCTTATCTTGGCTGCGCAGCAGAGATGCTCTCAACAAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuValLeuGlyCysGlnArgAspValLeuAsnLysGluMet 180
QY 541 CAGCAGAAATGGGATTTGGTATGTTAAATGCCAGCTAAATCCTGCAAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyTyrValIleAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 660
```

```
DB 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
QY 661 CCGTGGTTGGACAAATCACTAGTATTTGATGAGAAAGCAAAAGCTCCCAATGATGTGT 720
DB 221 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTAGTGCATGTTTGGTGGAGATCTCGGCTCGGCAACATGCTGCTGCTCATCATCAG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAGAAAGAAAGCACTACT 840
DB 261 LysArgMetAspMetSerLeuAsnGlyAlaTyrArgPheValLysGluLysArgProThr 280
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCCACTCTGAGATTAAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302

RESULT 6
US-10-343-357-7
/ Sequence 7, Application US/10343357
/ Publication No. US20040058341A1
/ GENERAL INFORMATION:
/ APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.Tom
/ APPLICANT: ELIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
/ APPLICANT: YAO, Monique G.; BUREFORD, Neil
/ APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
/ APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
/ APPLICANT: LEE, Ernestine A.; HAPALIA, April J.A.
/ APPLICANT: LU, Dying Aina M.; TRIBOLEY, Catherine M.
/ APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
/ APPLICANT: YUE, Henry; WARREN, Bridget A.
/ APPLICANT: NGUYEN, Daniel B.; CHAMLA, Nandinder K.
/ APPLICANT: KEARNEY, Liam
/ TITLE OF INVENTION: PROTEIN PHOSPHATASES
/ FILE REFERENCE: PI-0173 PCT
/ CURRENT APPLICATION NUMBER: US/10/343,357
/ PRIOR FILING DATE: 2003-01-28
/ PRIOR APPLICATION NUMBER: PCT/US01/23716
/ PRIOR FILING DATE: 2001-07-26
/ PRIOR APPLICATION NUMBER: US 60/221,679
/ PRIOR FILING DATE: 2000-07-28
/ PRIOR APPLICATION NUMBER: US 60/223,272
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: US 60/224,309
/ PRIOR FILING DATE: 2000-08-10
/ PRIOR APPLICATION NUMBER: US 60/226,728
/ PRIOR FILING DATE: 2000-08-18
/ PRIOR APPLICATION NUMBER: US 60/229,254
/ PRIOR FILING DATE: 2000-08-30
/ PRIOR APPLICATION NUMBER: US 60/231,366
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PERL Program
/ SEQ ID NO 7
/ LENGTH: 665
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CD1
/ US-10-343-357-7

Alignment Scores:
Pred. No.: 6, 03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
```

Query Match:

95.39%

Indels:

0

Gaps:

0

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-343-357-7 (1-665)

QY 1 ATGGCCCATGATGATGAACTCAATTTGTTAGAGGTTGGCTGCTGGAA 60
 DB 1 MetAlahisgluMetilegIythrGlnileValThrgluAglueValAlaleuLeuGlu 20
 QY 61 AGTGAACGGAAAAAGTGTCTAATGATGACCGGCAATTTGGAAATCAATACATCC 120
 DB 21 SerGlyThrgluValleuLeuileAspSerArgProPheValGluYrAsnThrSer 40
 QY 121 CACATTTGGAGCCATTAAATATCACTGCTCCAAAGCTTATGAAGGAGGTTGCAACG 180
 DB 41 HistileuGluAlaleAsnileAsnCysSerIlyseuMetIysArgIreuGlnGln 60
 QY 181 GACAAAGTGTAAATATACAGCTCATCCAGCTTACGCGAAACATTAAGGTTGACATGAT 240
 DB 61 AspIysValleuileThrgluLeuileGlnHisSerAlalyshIlyValAspIleAsp 80
 QY 241 TGCAGTCAGAGGTTGATGATTCAGTCAAAAGCTCCAGATGTTGCTCTCTCTTCA 300
 DB 81 CysSerGlnIysValIleValIyAspGlnSerSerGlnAspValAlaSerIleuSerSer 100
 QY 301 GACTGTTTCTCACTGACTTCTGGGTAACCTGGAGAGAGAGCTTCACTGTTCACTG 360
 DB 101 AspCysPheleuThrValleuLeuGlyIysleuGluIysSerPheAsnSerValHisIleu 120
 QY 361 CTTCAGAGTGGGTTGCTGAGTCTCTGTTGTTCCCTGGCCCTGTGAAAGAAATCC 420
 DB 121 LeuAlaGlyIyPheAlaIleuPheSerArgCysPheProGlyIysGlnIyIysSer 140
 QY 421 ACTGATGCTCACTGCAATTTCTCAGCTTGTCTTACCTGTCGCAACATTTGGGCCAAC 480
 DB 141 ThrleuValProThrCysIleSerGlnProCysleuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATTTCTCCCACTTATCTTCTGGCTGCGCAGGAGATGCTCTCAACAGAGCTGATA 540
 DB 161 ArgIleleuProAsnleuIyIreuGlyCysGlnArgAspValleuAsnIyIseuMet 180
 QY 541 CAGCAGATGGGATGTTGTTATGTTTAAATGACGCTATACCTGTCGCAACGCTGACTT 600
 DB 181 GlnGlnAsnGlyIleGlyIyValleuAsnAlaSerAsnThrCysProIysProAspPhe 200
 QY 601 ATCCCGCAGTCTCATTTCTGCGTGCCTGGATGATGACAGCTTTGTGAGAAATTTG 660
 DB 201 IleProGlnIserHisPheleuArgValIProValAsnAspSerPheCysGluIyIleleu 220
 QY 661 CCGTGGTTGGCAAAATCAGTATGTTTCAATGAGAAAGCAAAAGCTCCATGATGATGTT 720
 DB 221 ProThrleuAspIysSerValAspPheIleGlyIysAlaIysAlaSerAsnGlyCysVal 240
 QY 721 CTGATGCTGTTTATGCTGGATCTCCGCTCCGCAACATCCGCTATGCTTACATCATG 780
 DB 241 LeuValHisCysleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIyIleMet 260
 QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
 DB 261 LysArgMetAlaPheSerleuAspGluAlaIyArgPheValIyIseuIyysArgProThr 280
 QY 841 ATATCTCCAACTTCAATTTCTGGGCAACTCTGTCGATGAGAAAGAAAGATTAAGAAC 900
 DB 281 IleSerProAsnPheAsnPheleuGlyGlnleuLeuAspIyGluIyIysIleIysAsn 300
 QY 901 CAGACT 906
 DB 301 GlnThr 302

RESULT 7

US-10-377-072-26

Sequence 26 Application US/10377072

Publication No. US20040009501A1

/ GENERAL INFORMATION:
 / APPLICANT: Millennium Pharmaceuticals Inc.
 / APPLICANT: Curtis, Rory A.J.
 / APPLICANT: Logan, Thomas Joseph
 / APPLICANT: Glucksmann, Maria A.
 / APPLICANT: Meyers, Rachel E.
 / APPLICANT: Williamson, Mark J.
 / APPLICANT: Rudolph-Owen, Laura A.
 / APPLICANT: Chun, Myoung
 / APPLICANT: Tsai, Pong-Ying
 / TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117
 / TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
 / FILE REFERENCE: MP103-0180NMIM
 / CURRENT APPLICATION NUMBER: US/10/377,072
 / PRIOR APPLICATION NUMBER: 2003-02-27
 / PRIOR APPLICATION NUMBER: US 09/895,860
 / PRIOR FILING DATE: 2001-06-29
 / PRIOR APPLICATION NUMBER: US 60/215,370
 / PRIOR FILING DATE: 2000-06-29
 / PRIOR APPLICATION NUMBER: US 60/187,455
 / PRIOR FILING DATE: 2000-03-07
 / PRIOR APPLICATION NUMBER: US 09/843,297
 / PRIOR FILING DATE: 2001-04-25
 / PRIOR APPLICATION NUMBER: US 60/199,801
 / PRIOR FILING DATE: 2000-04-26
 / PRIOR APPLICATION NUMBER: US 09/861,801
 / PRIOR FILING DATE: 2001-05-21
 / PRIOR APPLICATION NUMBER: US 60/205,508
 / PRIOR FILING DATE: 2000-05-19
 / PRIOR APPLICATION NUMBER: US 09/816,494
 / PRIOR FILING DATE: 2001-03-23
 / PRIOR APPLICATION NUMBER: US 09/815,419
 / REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or PAM.
 / NUMBER OF SEQ ID NOS: 114
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO: 26
 / LENGTH: 665
 / TYPE: PRT
 / ORGANISM: Homo Sapiens
 US-10-377-072-26
 Alignment Scores:
 Pred. No.: 5,03e-152
 Score: 1552.00
 Percent Similarity: 99.67%
 Best Local Similarity: 99.34%
 Query Match: 95.39%
 DB: 15
 Gaps: 0
 US-10-029-345a-108_copy_538_1443 (1-906) x US-10-377-072-26 (1-665)
 QY 1 ATGGCCCATGATGATGAACTCAATTTGTTAGAGGTTGGCTGCTGGAA 60
 DB 1 MetAlahisgluMetilegIythrGlnileValThrgluAglueValAlaleuLeuGlu 20
 QY 61 AGTGAACGGAAAAAGTGTCTAATGATGACCGGCAATTTGGAAATCAATACATCC 120
 DB 21 SerGlyThrgluValleuLeuileAspSerArgProPheValGluYrAsnThrSer 40
 QY 121 CACATTTGGAGCCATTAAATATCACTGCTCCAAAGCTTATGAAGGAGGTTGCAACG 180
 DB 41 HistileuGluAlaleAsnileAsnCysSerIlyseuMetIysArgIreuGlnGln 60
 QY 181 GACAAAGTGTAAATATACAGGCTCATCCAGCTTACGCGAAACATTAAGGTTGACATGAT 240
 DB 61 AspIysValleuileThrgluLeuileGlnHisSerAlalyshIlyValAspIleAsp 80
 QY 241 TGCAGTCAGAGGTTGATGATTCAGTCAAAAGCTCCAGATGTTGCTCTCTCTTCA 300

US-10-029-345A-108_COPY_538_1443 (1-906) X US-10-257-026-2 (1-665)

1. TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR

```

; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 240
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-240

Alignment Scores:
Pred. No.: 6,03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-648-593-240 (1-665)

QY 1 ATGGCCCATGAGATGATTTGAACTCAAAATTGTTACTGAGAGGTGGTGGCTGCTGGA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTGCTGCTAATGATCCGGCCATTGGAATACATACATCC 120
Db 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
QY 121 CACATTTGGAGCCATTAAATATCACTGCTCAAGCTTATGAAGGAAAGTTGCAACG 180
Db 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleuMetLysArgLysLeuGln 60
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCAACATTAAGTTGACATTGAT 240
Db 61 AspLysValIleuIleThrGluLeuIleGlnHisSerAlaIleHisLysValAspIleAsp 80
QY 241 TGCAGTCGAGAGGTGTGATTTACGATCAAAAGCTCCCAAGATGGTGGCTCTCTCTCA 300
Db 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTTCTCACTGACTTCTGGGTAACGTGGAAGAGCTTCAACTCTGTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGTGGGTTGCTGAGTCTCTGTTGTTCCCTGGCTCTGTGGAAGAAATCC 420
Db 121 LeuAlaGlyValPheAlaGlnPheSerArgCysPheProGlyLeuGlyGluGlyLysSer 140
QY 421 ACTCTAGTCCCTACCTGATTTCTAGGCTTGCTTACTGCTGGCAACATTTGGGCCCAAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTCCCAATCTTATCTTGGCTGCGCAAGAGATGCTCTCAACAAGAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuMet 180
QY 541 CAGCAGATGGAGATTGTTATGTGTAAATGCCAGCATATCCGTGGCAAGAGCTTACTTT 600
Db 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTGGCGTGGCCCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
QY 661 CCGGTGTGGAACAATCATAGATTTTCAATTTGAGAAAGAAAGCTCCAAATGAGATGTT 720
Db 221 ProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 740
QY 721 CTAGTGCACTGTTTACTGAGATTCGCCGCTCCGACCAATGCTATTCGCTACATCATG 780

; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 247
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-247

Alignment Scores:
Pred. No.: 6,03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-648-593-247 (1-665)

QY 1 ATGGCCCATGAGATGATTTGAACTCAAAATTGTTACTGAGAGGTGGTGGCTGCTGGA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTGCTGCTAATGATCCGGCCATTGGAATACATACATCC 120
Db 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
QY 121 CACATTTGGAGCCATTAAATATCACTGCTCAAGCTTATGAAGGAAAGTTGCAACG 180
Db 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleuMetLysArgLysLeuGln 60
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCAACATTAAGTTGACATTGAT 240
Db 61 AspLysValIleuIleThrGluLeuIleGlnHisSerAlaIleHisLysValAspIleAsp 80
QY 241 TGCAGTCGAGAGGTGTGATTTACGATCAAAAGCTCCCAAGATGGTGGCTCTCTCTCA 300
Db 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTTCTCACTGACTTCTGGGTAACGTGGAAGAGCTTCAACTCTGTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGTGGGTTGCTGAGTCTCTGTTGTTCCCTGGCTCTGTGGAAGAAATCC 420

```

Db 121 LeuAlaGlyLpheaAlaGluPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTTGCAACATTTGGCCCAACC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCAGAGATGCTCTCAACAGAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuIyrLeuGlyCysGlnArgSerAlaLeuAsnIleGlyLysLeuMet 180
QY 541 CAGCAGAAATGGATGGTATGTTAATGACAGCTATACCTGTCACAAAGCTGACTT 600
Db 181 GlnGlnAsnIleGlyIyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCTCCGAGCTCATTTCTGCGTGCCTGCTGATGACAGCTTTTGTGGAATAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIyrIleLeu 220
QY 661 CCGTGGTGGACAAATCAGTATGATTCATTGAGAAAGCAAAAGCTTCACATGATGTT 720
Db 221 ProIrrPheAspLysSerValAspPheIleGlnIyrValAlaSerAsnIleCysVal 240
QY 721 CTAGTCACTGTTTAAAGTGGATCTCCGCTCCGACCACTGCTATGCTTACATCATG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIyrIleMet 260
QY 781 AAGAGATGACATGCTTCTTAAATGAACTTACAGATTTGTGAAAGAAAAAGCTTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaIyrArgPheValIyrGlnIyrAspProThr 280
QY 841 ATATCTCCCAATCTCATTTCTGGGCAACCTCCTGGACATGAGAAAGATTAAGAAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlyGlnIleLeuAspIyrGlnIyrLysIleLysAsn 300
QY 901 CAGACT 906
Db 301 GlnThr 302

RESULT 11
US-10-072-012-679
; Sequence 679, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Speyck, Kimberly
; APPLICANT: Zehrusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Beha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimyr Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Molenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514

QY 121 LeuAlaGlyLpheaAlaGluPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCAGAGATGCTCTCAACAGAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuIyrLeuGlyCysGlnArgSerAlaLeuAsnIleGlyLysLeuMet 180
QY 541 CAGCAGAAATGGATGGTATGTTAATGACAGCTATACCTGTCACAAAGCTGACTT 600
Db 181 GlnGlnAsnIleGlyIyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCTCCGAGCTCATTTCTGCGTGCCTGCTGATGACAGCTTTTGTGGAATAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIyrIleLeu 220
QY 661 CCGTGGTGGACAAATCAGTATGATTCATTGAGAAAGCAAAAGCTTCACATGATGTT 720
Db 221 ProIrrPheAspLysSerValAspPheIleGlnIyrValAlaSerAsnIleCysVal 240
QY 721 CTAGTCACTGTTTAAAGTGGATCTCCGCTCCGACCACTGCTATGCTTACATCATG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIyrIleMet 260
QY 781 AAGAGATGACATGCTTCTTAAATGAACTTACAGATTTGTGAAAGAAAAAGCTTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaIyrArgPheValIyrGlnIyrAspProThr 280
QY 841 ATATCTCCCAATCTCATTTCTGGGCAACCTCCTGGACATGAGAAAGATTAAGAAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlyGlnIleLeuAspIyrGlnIyrLysIleLysAsn 300
QY 901 CAGACT 906
Db 301 GlnThr 302

US-10-029-345A-108_COPY_538_1443 (1-906) x US-10-072-012-679 (1-690)

QY 1 ATGGCCCATGAGATGATGAACTCAATTTGTTACTGAGAGTGGTGGCTGTGTA 60
Db 26 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuGln 45
QY 61 AGTGAACGAAAAAGTGTGCTAATGATGAGCCGCAATTTGTGGAATACATATCC 120
Db 46 SerGlyThrGlnIyrValLeuLeuIleAspSerArgProPheValGlnIyrAsnThrSer 65
QY 121 CACATTTGGAAGCCCTTAATATCACTGCTCCAACTTATGAAGGAGCTTGCAACG 180
Db 66 HisIleLeuGlnAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 85
QY 181 GACAAAGTGTAAATTCAGAGCTCATCCAGCAATTCAGCAAAACATAGGTGATGAT 240
Db 86 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaIyrHisIyrValAspIleAsp 105
QY 241 TGCAGTCAGAAAGTGTGATTTAAGTCAAAAGCTCCCAAGATGTTGCTCTCTTCA 300
Db 106 CysSerGlnIyrValIyrValIyrValIyrValIyrValIyrValIyrValIyrVal 125
QY 301 GACTGTTTCTCACTGATCTTCTGCGTAACTGAGAAAGCTTCAACTGTTCACTG 360
Db 126 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnIyrSerPheAsnSerValHisLeu 145
QY 361 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTGTTTCCCTGCGCTGTGGAAGAAATCC 420
Db 146 LeuAlaGlyLpheaAlaGluPheSerArgCysPheProGlyLeuGlyGlyLysSer 165
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTTGCAACATTTGGCCCAACC 480
Db 166 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 185
QY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCAGAGATGCTCTCAACAGAGAGCTGATA 540
Db 186 ArgIleLeuProAsnLeuIyrLeuGlyCysGlnArgSerAlaLeuAsnIleGlyLysLeuMet 205
QY 541 CAGCAGAAATGGATGGTATGTTAATGACAGCTATACCTGTCACAAAGCTGACTT 600

```
Db 206 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProIysProAspPhe 225
QY 601 ATCCCCAGATCTCATTTCTCGCTGTCCTGCTGAATGACAGCTTTTGTGAAATTTTG 660
Db 226 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleLeu 245
QY 661 CCGTGGTTGGACAAATCAGTAGATTTCATTTGAGAAAGCAAAAGCTTCCAAATGATGTGT 720
Db 246 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 265
QY 721 CTAGTGCACTGTTTGGCTGGATCTCCGCTCCGCAACATCCCTATCCGCTACATCATG 780
Db 266 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaValIleMet 285
QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTGTGAAAGAAAAAGACTACT 840
Db 286 LysArgMetAspMetSerLeuAspGluAlaIleArgPheValLysGluLysArgProThr 305
QY 841 ATATCTCCAACTCATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
Db 306 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 925
QY 901 CAGACT 906
Db 326 GlnThr 327

RESULT 12
US-10-072-012-703
; Sequence 703, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patcurajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esba
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosee, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072, 012
; PRIOR APPLICATION NUMBER: 2002-01-31
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265, 514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265, 517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265, 412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265, 395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266, 406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266, 767
; PRIOR FILING DATE: 2001-02-05
```

```
; PRIOR APPLICATION NUMBER: 60/267, 057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266, 975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267, 459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 703
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-703

Alignment Scores:
Pred. No.: 6,09e-152
Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Query Match: 95.39%
DB: 12
Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-072-012-703 (1-690)
QY 1 ATGGCCCATGAGATGATTGGAACTCAATTGTTACTGAGAGCTTGCTGCTGGAA 60
Db 26 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 45
QY 61 AGTGAACCGAAAAAGTCTGCTAATTGATGCGCGCATTTGTGAAATACATATCATC 120
Db 46 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluValAsnThrSer 65
QY 121 CAATTTTGAAGCCATTAATATCAACTGCTCCAGCTTATGAGCGAAGTGGCAACG 180
Db 66 HisIleGluAlaIleHisIleAsnLysCysSerLysLeuMetLysArgLysGluInd 85
QY 181 GACAAAGCTTAATTAACAGCTCATTCAGCATTCAGCGAAACATAGAGTTGACATTGAT 240
Db 86 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspLeu 105
QY 241 TGCAGTCAGAGGTGATGATTACGATCAAGCTCCCAAGATGTTGCTCTCTTCA 300
Db 106 CysSerGlnLysValValValIleArgAsnSerSerGlnAspValAlaSerLeuSer 125
QY 301 GACTGTTTCTCATCTGATCTTCTGAGTAACTGAGAGAGCTTCAACTGTTCACTG 360
Db 126 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 145
QY 361 CTTCAGGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATCC 420
Db 146 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlySer 165
QY 421 ACTCTAGTCCCACTGATTTCTGAGCTTCTTACCTTCTTACCTTCCAACTTGGGCAACC 480
Db 166 ThrIleValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 185
QY 481 CCAATTTTCCCAATCTTATCTTGCTGCGCAGAGAGTCTCTCAACAGAGAGCTGATA 540
Db 186 ArgIleLeuProAsnLeuIleLysGlyCysGlnAspValLeuAsnLysGluLeuMet 205
QY 541 CAGCAGAAATGGGATTTGTTTATGTTAATGCCAGCTATACCTGCAAGAGCTGATTT 600
Db 206 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 225
QY 601 ATCCCCAGATCTCATTTCTCGCTGTCCTGCTGAATGACAGCTTTTGTGAAATTTTG 660
Db 226 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleLeu 245
QY 661 CCGTGGTTGGACAAATCAGTAGATTTCATTTGAGAAAGCAAAAGCTTCCAAATGATGTGT 720
Db 246 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 265
```

```

Oy      721 TAGAGCACTGTTTAGTCGGAGATCCCGCTCGGCACCATGGCTTATGCCCTTAACATCANG   780
       |||||
Db      266 LeuValHisCySLeuAlaIleYllSerArgSerAlaThrIleAlaIleAlaTyrlleMet   285
       |||||
Oy      781 AAGAAGATGACATGCTCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAAAGACTTACT   840
Db      286 LysArgMetLapMetSerLeuAspGluAlaTyTrArgPheValLlysGluLysArgProThr   305
Oy      841 ATATTCCAAACCTTGCAATTTTTCTGGGCCCAACTCTTGAGCATATGAGAAAGAATTAAAGAA   900
Db      306 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTy-GlnLysLysIleLysAsn   325
Oy      901 CAGACT   906
       |||||
Db      326 GlnThr   327

RESULT 13
US-10-425-114-54204
; Sequence 54204, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54204
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-028-H6_FLI pep
; US-10-425-114-54204
```

Alignment Scores:	
Prod. No.:	6, 09e-152
Score:	1552.00
Percent Similarity:	99.67%
Best Local Similarity:	99.34%
Query Match:	95.39%
DB:	12
	Gaps: 0
	Length: 690
	Matches: 300
	Conservative: 1
	Mismatches: 1
	Indels: 0

US-10-029-345A-108_COPY_538_1443 (1-906) X US-10-425-114-54204 (1-690)

Qy	1	ATGGCCCATGAGATGATTTGGAAATCGAAATTTGTTACTGAGAGAGTTGTGGCTCTGTGGAA	60
Db	26	MetAlaHisGluIuectIleGlyThrGlnIleValThrGluArgLeuValAlaIleuLeuGlu	45
Qy	61	AGTGAACGCAAAAAGTGTGCTGATTTAGATTCACCGGCGCTTGTGTGAATTAACAATCATCC	120
Db	46	SecGlyThrClnIuyluylValIleuLeuIleAspSerAsnGlyProPheValGlnTybAsnThrSer	65
Qy	121	CACATTTGGAGCCATTAATATGAACTGCTCCGAGCTTATGAGACGGAAGTTGCACAG	180
Db	66	HisIleLeuGlnIuAlaIleAsnIleAsnIleAsnCySerIuyluMetIuyluArgIleGlnGlnGln	85
Qy	181	GACAAAGTGTAAATTTACAGAGCTCATCCAGCAATTCAGGAAACATTAAGTTGACACTGAT	240
Db	86	AspIuyluValIleuIleThrGlnIleuIleGlnHisSerAlaIuyluValAlaAspIleAsp	105
Qy	241	TGCAGTCGAGAGGTTGTAGTTTACGATCAAGCTCCAGAGATGTCCTCTCTCTTCA	300
Db	106	CysSecGlnIuyluylValIuyluValIuyluArgIuyluInsSerGlnAspValAlaSerIleuSerSer	125
Qy	301	GACTGTTTTCACCTGTACTTCTGGGTAAATCGAGAGAGCTTCAACTCTGTTCACTTG	360

Db	126	AspGyPheIeuThrValIleuLeuGlyLysIleuGluLysSerPheAsnSerValHisIeu	145
Qy	361	CTTGCAAGTGGGTGTTGCTGAGTCTCTGTTGTTTCCCTGGGCTCTGTGAAGAAATCC	420
Db	146	LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlySerGluLysSer	165
Qy	421	ACTAGTACCCGACCTGCAATCTTCTGAGCTTCCTTACCTGTTGGCCAACTTGGGCCAAC	480
Db	166	ThrIleuValProThrCysAlIleSerGlnProCysIleuProValAlaAsnIleGlyProThr	185
Qy	481	CGAATTCCTCCCAATCTTTATCTTGAGCTGCGCCAGCAGATGTCCTCAACAAGAGCTGATA	540
Db	186	ArgIleIeuProAsnLeuTyrIleuGlyCysGlnArgAspValIleuAsnLysGluIleuMet	205
Qy	541	CAGCAGAATGGATGTTGTTATGTTAAATGCCAGCTAATCCTGTCCAAAGCTGACTTT	600
Db	206	GlnGlnAsnGlyIleGlyTyrValIleuAsnAlaSerAsnThrCysProLysProAspPhe	225
Qy	601	ATCCCGAGTCTCATTTCCGCGTGTGCCTGTGATGACAGCTTTTGTGAGAAAAATTTTG	660
Db	226	IleProGluSerHisPheIleuArgValProValAsnAspSerPheCysGluLysIleIeu	245
Qy	661	CCGTGGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCATGGATGTGTT	720
Db	246	ProTyrIeuAspLysSerValAspPheIleGlyLysValAlaLysAlaSerAsnGlyCysVal	265
Qy	721	CTAGTGCACCTGTTAGCTGGGATCTTCCGGCTCCGCCACCATGTGCTATGCCCTACATCANG	780
Db	266	LeuValHisCysIleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet	285
Qy	781	AAGAGATGAGCATGTCTTTAGATAGAGCTTCACAGATTGTGAAGAAAGAAAGACTTACT	840
Db	286	LysArgMetAspMetSerIleuAspGluAlaTyrArgPheValLysGluLysArgProThr	305
Qy	841	ATATCTCCAAACTTCATTTTCTGGGCCCACTCTGAGCTATGAGAAAGAGATTAAAGAAC	900
Db	306	IleSerProAsnPheAsnIleuGlyIleIleuLeuAspTyrGluLysIleLysAsn	325
Qy	901	CAGACT 906	
Db	326	GlnThr 327	

RESULT 14
US-10-072-012-681
; Sequence 681, Application US/10072012
; Publication No. US20040003493A1
; GENERAL INFORMATION:

APPLICANT: Spytek, Kimberly
APPLICANT: Zernhusen, Bryan
APPLICANT: Patcurajan, Meera
APPLICANT: Shmkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimirt Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Putrak, Katarzyna
APPLICANT: Grosee, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012

```

; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: Patentln Ver. 2.1
; SEO ID NO: 681
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-681

```

```

Alignment Scores:
Pred. No.: 3,21e-151 Length: 665
Score: 1545.00 Matches: 299
Percent Similarity: 99.34% Conservative: 1
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 94.96% Indels: 0
DB: 12 Gaps: 0

```

```

US-10-029-345A-108_COPY_538_1443 (1-906) x US-10-072-012-681 (1-665)

```

```

QY 1 ATGGCCCATGAGTATGATGTTGAACCTCAATTTGTTACTGAGAGGTTGGTCTCTGCGAA 60
DB 1 MetAlahIsgIumetIlegIyThrgInIleValThrgIuAgIueValAlaLeuIeuglu 20
QY 61 AGTGAACGGAAGGTTGCTGCAATTTGATGACCGGCAATTTGGATATACATATACATCC 120
DB 21 SerGIyThrcIuIyValIeuIleAspSerArGpPheValGIuTyAenThSer 40
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAAAGTTTATGAGCGAAGTTGCAACAG 180
DB 41 HisIleIeugluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGGAAACATTAAGTTGACATTGAT 240
DB 61 AspIyValIeuIleThrcIuIeugIleGInIleGInIleGInIleGInIleGInIleGIn 80
QY 241 TGCAGTCGAGAGGTTGATGTTAGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB 81 CysSerGIuIyValIeValIyAspGIuIeSerSerIeAspValAlaSerIeuSer 100
QY 301 GACTGTTTTCACGTCATCTCTGCTGTTTCCGAGGCTCTGAGAGGAAATCC 420
DB 101 AspCysPheIeuThrcIuIeugIyIyValIeugIuIySerPheAsnSerValIHisIeu 120
QY 361 CTTCGAGGTGGGTTTGGTGAATCTCTGTTGTTTCCGAGGCTCTGAGAGGAAATCC 420
DB 121 LeuAlaGIyGIyPheAlaGIuIeSerArGpCysPheProGIyIeucySerIuIySer 140
QY 421 ACTGTAGTCCCTACCTGATTTCTCAGCCTTGTTTCTGTTTCCGAGGCTCTGAGAGG 480
DB 141 ThrIeuValIProThrcIyIleSerGIuIeProCysIeuProValAlaAsnIleGIyProThr 160
QY 481 CGAATTTCCCAATCTTATCTTGCTGCGACGAGATGTCCTCAAGAAGACTGATA 540

```

```

DB 161 LeuIleIeuProAsnIeuIyIeugIyIySerGIuIeAspValIleuAsnIySGIuIeMet 180
QY 541 CAGCAGAAATGAGATGTTGTTATGTTGTTAAATCCCGCTATACCTGTCGAAAGCTGACTTT 600
DB 181 GIuGIuAsnGIyIleGIyIyValIeAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 200
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 201 IleProGIuIeSerIeAspIeAsnIleValIProValIAsnAspSerIeIeGIyIleIeu 220
QY 661 CCGTGGTTGACAATATCATGATGATTTTCATTTGAGAGGAAAGGCTCCATGATGATGTT 720
DB 221 ProIleuAspIySerValIAspIeIleGIuIyAlaIyAlaIleAsnIleGIyValI 240
QY 721 CTAGTGACCTGTTTATGCTGGGATCTCCCGCTCCGACCATGCTATGCGCTTACATCANG 780
DB 241 LeuValIhScyIeuAlaGIyIleSerArGSerAlaThrIleAlaIleAlaIyIleIeMet 260
QY 781 AAGAGATGACATGCTTTTATGATGAGCTTACAGATTTGTGAAAGGAAAGGAAAGGACTTACT 840
DB 261 LysArGIeAspIeSerIeAspGIuIaIyIyArGpIeValIySGIuIyArGProThr 280
QY 841 ATATCCCAACTTCAATTTTCTGCGCAACTCTGAGCTATGAGAGGATTTAAGAC 900
DB 281 IleSerProAsnIeAsnIeIeugIyGIuIeIeuAspIyGIuIyGIyIleIyAsn 300
QY 901 CAGACT 906
DB 301 GIuThr 302

```

```

RESULT 15

```

```

; Sequence 2312, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHITO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patentln Ver. 2.1
; SEO ID NO: 2312
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2312

```

```

Alignment Scores:
Pred. No.: 3,21e-151 Length: 665
Score: 1545.00 Matches: 299

```

Percent Similarity: 99.34%
Best Local Similarity: 99.01%
Query Match: 94.96%
DB: 15
Conservative: 1
Matches: 2
Indels: 0
Gaps: 0

Job time : 51.0982 secs

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-094-749-2312 (1-665)

```
QY 1 ATGGCCATGATGATTTGGAACCTCAAAATGTTACTGAGAGGTTGGTGGCTGCTGGAA 60
DB 1 MetAlaHisGlnMetCileGlyThrGlnIleValThrGlnArgLeuValAlaLeuGln 20
QY 61 AGTGAACGGAAAAAGTGTCTAATTGATAGCCGGCCATTGTGGATACAATACATCC 120
DB 21 SerGlyThrGlnValLeuLeuIleAspSerArgProPheValGluTyraThrSer 40
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAACTTATGAAGCAAGTTGCAACAG 180
DB 41 HisIleuGlnIleAlaIleAsnIleAsnCySerIleuMetLysArgLeuGln 60
QY 181 GACAAAGTGTATTACAGAGCTCATCCAGCATTCAGCAACATAAGTTGACATTGAT 240
DB 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCAGTCAGAAAGGTGATGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTTCA 300
DB 81 CysSerGlnLysValValValTyraAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTTCACCTGACTGCTTGGGGTAACTGGAGAAAGCTTCAACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGGTGGTGGTGGTGGTCTCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlyLysSer 140
QY 421 ACTCTAGTCCCTACCTGCACTTCTCAGCTTGTGCTTACCTGTTGCCAACATTTGGGCCAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTCCTCCCAATCTTTATCTTGGCTGCCAGGAGATGCTCTCAACAAAGAGCTGATA 540
DB 161 LeuIleLeuProAsnLeuTyrlLeuGlyCysGlnArgAspValLeuAsnLysGlnLeuMet 180
QY 541 CAGCAGAAATGGATGGTGTATGTTAAATGCCAGCATATACCTGTCCAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyTyrlValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
QY 661 CCGTGTGGACAATGATGATGATTTGAGAAAGCAAAAGCTTCCAAATGATGTGTT 720
DB 221 ProThrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTAGTGACACTGTTAGCTGGGATCTCCCGCTCCGCCACCATGCTATCGCTTACATCAG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlLeuMet 260
QY 781 AAGAGATGACATGCTCTTATGATGAAGCTTACAGATTGTGAAGAAAAAGACTTACT 840
DB 261 LysArgMetCysPheSerLeuAspGlnAlaTyrlArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCCAACTCTTGACTATGAGAAAGATTAAAGAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrlGlnLysLysIleLysAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302
```

Search completed: June 21, 2004, 13:05:41

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:26 ; Search time 11.5553 Seconds

(without alignments)
8095.510 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 1627
Sequence: 1 atggcccatgagatgatgttg.....agaagataagaaccagact 906

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10029345/runat_21062004_122816_4233/app_query.fasta_1.3278
-DB=Issued Patents AA -OPMT=fastaan -SUFPIX=rai -MINMATCH=0.1 -LOOPLC=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345_QCGN_1_1_45 @runat_21062004_122816_4233 -NCPU=6
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARM TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA.*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep.*
4: /cgn2_6/prodata/2/1aa/5B_COMB.pep.*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1552	95.4	665	4	US-09-816-494-2
2	654.5	40.2	170	4	US-09-844-716-14
3	654.5	40.2	170	4	US-09-557-921-15
4	654.5	40.2	170	4	US-09-564-357-17
5	654.5	40.2	170	4	US-09-619-380-16
6	469	28.8	482	4	US-09-557-921-2
7	427	26.2	394	4	US-08-530-290-23
8	427	26.2	394	4	US-09-702-705-805
9	427	26.2	394	4	US-09-702-705-827
10	427	26.2	394	4	US-09-736-457-805
11	427	26.2	394	4	US-09-736-457-827
12	427	26.2	394	4	US-09-614-124B-805

13	427	26.2	394	4	US-09-614-124B-827	Sequence 827, App
14	427	26.2	394	4	US-09-671-325-805	Sequence 805, App
15	427	26.2	394	4	US-09-671-325-827	Sequence 827, App
16	427	26.2	394	4	US-09-589-184-805	Sequence 805, App
17	427	26.2	394	4	US-09-589-184-827	Sequence 827, App
18	425	26.1	394	4	US-09-371-671B-11	Sequence 11, App
19	425	26.1	394	4	US-08-990-379-6	Sequence 6, App
20	423.5	26.0	395	2	US-08-990-379-5	Sequence 5, App
21	415	25.5	367	2	US-08-530-290-24	Sequence 24, App
22	409.5	25.2	313	2	US-08-990-379-7	Sequence 7, App
23	409	25.1	314	3	US-09-164-193-22	Sequence 22, App
24	409	25.1	314	4	US-09-921-448A-22	Sequence 22, App
25	377.5	23.2	393	2	US-08-990-379-4	Sequence 4, App
26	357	21.9	302	4	US-09-702-705-806	Sequence 806, App
27	357	21.9	302	4	US-09-736-457-805	Sequence 806, App
28	357	21.9	302	4	US-09-614-124B-806	Sequence 806, App
29	357	21.9	302	4	US-09-671-325-806	Sequence 806, App
30	357	21.9	302	4	US-09-589-184-806	Sequence 806, App
31	349.5	21.5	397	2	US-08-990-379-8	Sequence 8, App
32	346.5	21.3	168	4	US-09-544-716-13	Sequence 13, App
33	346.5	21.3	168	4	US-09-557-921-13	Sequence 13, App
34	346.5	21.3	168	4	US-09-564-357-16	Sequence 16, App
35	346.5	21.3	168	4	US-09-619-380-15	Sequence 15, App
36	338.5	20.8	170	4	US-09-544-716-12	Sequence 12, App
37	338.5	20.8	170	4	US-09-557-921-12	Sequence 12, App
38	338.5	20.8	170	4	US-09-564-357-15	Sequence 15, App
39	338.5	20.8	170	4	US-09-619-380-14	Sequence 14, App
40	319	19.6	169	4	US-09-544-716-16	Sequence 16, App
41	319	19.6	169	4	US-09-557-921-17	Sequence 17, App
42	319	19.6	169	4	US-09-564-357-19	Sequence 19, App
43	319	19.6	169	4	US-09-619-380-18	Sequence 18, App
44	312	19.2	168	4	US-09-544-716-15	Sequence 15, App
45	312	19.2	168	4	US-09-557-921-16	Sequence 16, App

ALIGNMENTS

RESULT 1
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816, 494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191, 858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FaetsEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-494-2

Alignment Scores:
Pred. No.: 3.93e-181
Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Query Match: 95.39%
Gaps: 4
Indels: 0
Matches: 300
Conservative: 1
Mismatch: 1
Indel: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-816-494-2 (1-665)

QY 1 ATGGCCCATGAGATGATTTGGAAGTCAATTTGTTACTGAGAGCTGTGCTGCTGCGAA 60
DB 1 Mettllahlgllwctlllellythrglnllevalthghlaryglvvalalaleuileuglu 20
QY 61 AGTGAACGGAAGAAAGTGTGCTTAATTGATAGCCGSCCATTTGTGGAATACATACATCC 120

```

Db      21  SerGIyThrgIuIySvalIeuIleuIleAspSerArgProPheValGIuIyRasnThrSer 40
Qy      121  CACATTTTGGAGCCATTAAATATCAACTGCTCCAGCTTATTAAGCGAAAGGTTGCAACG 180
Db      41  HistIleuGlnAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 180
Qy      181  GACAAAGTGTAAATTAACAGAGCTCATCCGACCTTCAAGCAAAATAGTGAATGATCATTTAT 240
Db      61  AspIySValIleuIleThrgIuIleuIleGlnHisSerAlaIySHISlySValAspIleAsp 80
Qy      241  TGCAGTCAGAGGTTTGTATTAACATCAAAAGCTCCCAAGATGTTGCTCTCTCTCTTCA 300
Db      81  CysSerGIuIySValIyValIyRaspGIuIySerGIuIySerGIuIySerGIuIySerGIuIy 100
Qy      301  GACTGTTTCTCACTGTAATCTTGGGTAAATCGAGAAAGCTTCAACTGTTCACCTG 360
Db      101  AspCysPheIeuThrgValIleuIleuIySValIleuIySerPheAsnSerValHisIleu 120
Qy      361  CTTGCAAGTGGGTTTGTGAGTTCCTGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db      121  LeuAlaGIyGIyPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 140
Qy      421  ACTCTAGCTTCTCACTGTAATCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db      141  ThrIleuValProThrCysIleSerGIuIySerGIuIySerGIuIySerGIuIySerGIuIy 160
Qy      481  GAAATTTCTTCCCAATCTTATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db      161  ArgIleuProAsnIleuIyRleuGIyCysGIuIyAspValIleuAsnIySGIuIleuIleu 180
Qy      541  CACAGATTTGGAGTGGTGTATGTATTAATGCAAGCTTATCACTGCTCAAGCTGCTGCTGCT 600
Db      181  GlnGlnAsnGIyIleIyIyValIleuAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 620
Qy      601  ATCCCGAGCTGCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db      201  IleProGIuIySerHisPheIleuArgValProValAsnAspSerPheCysGIuIyIleu 220
Qy      661  CCGTGGTGGAGCAAAATCGATTTGATTAATGAGAAAGCAAAAGCTTCCCAATGATGCTTT 720
Db      221  ProIleuAspIySValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 740
Qy      721  CTAGTGCCTGTTTAGCTGAGATCTCCGCTCCGACATGCTATGCTATGCTATGCTATGCTAT 780
Db      241  LeuValHisCysLeuAlaGIyIleSerArgSerAlaIleAlaIleAlaIleAlaIleAlaIle 260
Qy      781  AAGAGATGAGACATGCTTATGATGAGCTTACAGATTTGAGAAAGAAAGAAAGCTTACT 840
Db      261  LysArgMetAspMetSerIleuAspGlnAlaIyRArgPheValIySGIuIySArgProThr 280
Qy      841  ATATCTCCAAATCTCAATTTTCTGGGCAACTCCGAGCTATGAGAAAGAAAGATTAAGAAC 900
Db      281  IleSerProAsnPheAsnIleuGIyIleuIleuAspIyRGIuIySlySlyIleIySAsn 920
Qy      901  CAGACT 906
Db      301  GlnThr 302

```

```

; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-544-716-14

Alignment Scores:
Pred. No.: 2,226-71
Score: 654.50
Percent Similarity: 88.62%
Best Local Similarity: 72.46%
Query Match: 40.23%
DB: 4
Gaps: 1

US-10-029-345a-108_copy_538_1443 (1-906) x US-09-544-716-14 (1-170)
Qy      400  GGCCTGTGTAAGAGAAA---TCACCTAGTCCCTACCTGCAATTTCTCAGCTTGTCTTA 456
Db      1  GIIyIeucysGIuGIyIySProAlaAlaIleuIleuPrometSerIleuSerIleuProCysIleu 20
Qy      457  CCTGTGCAACATTTGGGCAACCCGCAATTTCCCAATCTTATCTTATCTTGGCTGCGACGA 516
Db      21  ProValProSerValGIyIleuThrgIleuIleuIleuIleuIleuIleuIleuIleuIleu 40
Qy      517  GATGCTCTCAACAGAGCTGATACAGAGAAATGGGATTTGTTATGTTTAAATGCCAGC 576
Db      41  AspValIleuAsnIySAspIleuMetThrgIleuIleuIleuIleuIleuIleuIleuIleu 60
Qy      577  TATACCTGTCCAAAGCTGATTTATCCCGAGTCTCAATTTCCGCTGCTGCTGCTGCTGCTG 636
Db      61  AsnSerCysPheIySProAspPheIleCysGIuIySerArgPheMetArgValProIleuAsn 80
Qy      637  GACAGCTTTTGAAGAAATTTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Db      81  AspAsnIyRGIySGIuIyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 100
Qy      697  GCAAAAGCTCCCAATGATGTGTTCTAGTCACTGTTTGGAGTCCGAGATCCCGCTCCGCC 756
Db      101  AlaIySleuSerSerCysGIuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 120
Qy      757  ACCATGCTATGCTTATCATATGAGAGAGATGACATGCTTTAGATGAGATTAAGAGA 816
Db      121  ThrIleAlaIleAlaIyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 140
Qy      817  TTGTGAAAGAAAGAAAGCTTCAATATCTCCAAATCTCAATTTTCTGGGCAACTCTCG 876
Db      141  PheValIySAspArgArgProSerIleSerProAsnIleuIleuIleuIleuIleuIleu 160
Qy      877  GACTATGAGAGAGATTAAG 897
Db      161  GlnIyRGIuIyArgThrgIleuIyS 167

RESULT 3
US-09-557-921-15
; Sequence 15, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Lucie, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT APPLICATION NUMBER: US/09/557,921
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-921-15

Alignment Scores:
Pred. No.: 2,226-71
Score: 654.50
Length: 170
Matches: 121
Conservative: 121
Mismatch: 18
Indels: 1
Gaps: 1

```

Percent Similarity: 88.62% Conservative: 27
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 40.23% Indels: 1
Gaps: 1
US-10-029-345a-108_COPY_538_1443 (1-906) x US-09-557-921-15 (1-170)

QY 400 GGGCTCTGTGAAGAAAA--TCACCTTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTA 456
DB 1 G1yleuCyglucllylsProAlaAlaLeuLeuProMeSerleuSerInProCySleu 20
QY 457 CCTGTGGCAACATGGGGCCAAACCGCAATTTCTCCCAATTTTATCTTGCTGCCAGCA 516
DB 21 ProValProSerValG1yleuThrArgIleLeuProH1leuYrleuGlySerGlnlys 40
QY 517 GATGTCTCAACAAGAGCTGATATACAGCAAGATGGATGGTATGTGTTAAATGCCAGC 576
DB 41 AspValLeuAsnLysAspLeuMetThrGlnAsnGlyIleSerYrValLeuAsnAlaSer 60
QY 577 TATACCTGTCAAAAGCTGACTTATCCCGAGTCTCATTTCCGTGCTGCTGCAAT 636
DB 61 AsnSerCySProLysProAspPheIleCySgluSerArgPheMetArgValProIleAsn 80
QY 637 GACAGCTTTGTGAGAAAAATTTGGCGGTGGTGGACAAATCAGATGATTCAGGAAA 696
DB 81 AspAsnTyCySgluLysLeuLeuProTyrLeuAspLysSerIleGlnPheIleAspLys 100
QY 697 GCAAAAGCTCCAAATGATGTGTCTAGTCACTGTTAGCTGGATCTCCGCTCGCC 756
DB 101 AlAlyLeuSerSerCySgluValIleValHisCySleuAlaGlyIleSerArgSerAla 120
QY 757 ACCATGCTATGCGCTACATCATGAGAGATGACATGCTTTAGATGAAGCTTACAGA 816
DB 121 ThrIleAlaIleAlaTyrlleMetLysThrMetGlyMetSerSerAspAlaTyArg 140
QY 817 TTTGTGAAGAAAAAAGCTTACTATATCTCAAACTTCAATTTTCTGGGCCAACTCTG 876
DB 141 PheValLysAspArgArgProSerIleSerProAsnPheAsnIleuGlyInLeuLeu 160
QY 877 GACTATGAGAGAAGATTAG 897
DB 161 GluTyrgluArgThrLeuLys 167

RESULT 4
US-09-564-357-17
Sequence 17, Application US/09564357
Patent No. 6645753
GENERAL INFORMATION:
APPLICANT: LucHe, Ralf M.
TITLE OF INVENTION: DSP-5 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.413
CURRENT APPLICATION NUMBER: US/09/564,357
CURRENT FILING DATE: 2000-04-24
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 170
TYPE: PRT
ORGANISM: Homo sapiens
US-09-564-357-17

Alignment Scores:
Pred. No.: 2,22e-71 Length: 170
Score: 654.50 Matches: 121
Percent Similarity: 88.62% Conserved: 27
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 40.23% Indels: 1
Gaps: 1
US-10-029-345a-108_COPY_538_1443 (1-906) x US-09-564-357-17 (1-170)

QY 400 GGGCTCTGTGAAGAAAA--TCACCTTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTA 456

DB 1 G1yleuCyglucllylsProAlaAlaLeuLeuProMeSerleuSerInProCySleu 20
QY 457 CCTGTGGCAACATGGGGCCAAACCGCAATTTCTCCCAATTTTATCTTGCTGCCAGCA 516
DB 21 ProValProSerValG1yleuThrArgIleLeuProH1leuYrleuGlySerGlnlys 40
QY 517 GATGTCTCAACAAGAGCTGATATACAGCAAGATGGATGGTATGTGTTAAATGCCAGC 576
DB 41 AspValLeuAsnLysAspLeuMetThrGlnAsnGlyIleSerYrValLeuAsnAlaSer 60
QY 577 TATACCTGTCAAAAGCTGACTTATCCCGAGTCTCATTTCCGTGCTGCTGCAAT 636
DB 61 AsnSerCySProLysProAspPheIleCySgluSerArgPheMetArgValProIleAsn 80
QY 637 GACAGCTTTGTGAGAAAAATTTGGCGGTGGTGGACAAATCAGATGATTCAGGAAA 696
DB 81 AspAsnTyCySgluLysLeuLeuProTyrLeuAspLysSerIleGlnPheIleAspLys 100
QY 697 GCAAAAGCTCCAAATGATGTGTCTAGTCACTGTTAGCTGGATCTCCGCTCGCC 756
DB 101 AlAlyLeuSerSerCySgluValIleValHisCySleuAlaGlyIleSerArgSerAla 120
QY 757 ACCATGCTATGCGCTACATCATGAGAGATGACATGCTTTAGATGAAGCTTACAGA 816
DB 121 ThrIleAlaIleAlaTyrlleMetLysThrMetGlyMetSerSerAspAlaTyArg 140
QY 817 TTTGTGAAGAAAAAAGCTTACTATATCTCAAACTTCAATTTTCTGGGCCAACTCTG 876
DB 141 PheValLysAspArgArgProSerIleSerProAsnPheAsnIleuGlyInLeuLeu 160
QY 877 GACTATGAGAGAAGATTAG 897
DB 161 GluTyrgluArgThrLeuLys 167

RESULT 5
US-09-619-380-16
Sequence 16, Application US/09619380
Patent No. 6643391
GENERAL INFORMATION:
APPLICANT: LucHe, Ralf M.
TITLE OF INVENTION: DSP-11 DUAL SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.418
CURRENT APPLICATION NUMBER: US/09/619,380
CURRENT FILING DATE: 2000-07-19
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 170
TYPE: PRT
ORGANISM: Homo sapiens
US-09-619-380-16

Alignment Scores:
Pred. No.: 2,22e-71 Length: 170
Score: 654.50 Matches: 121
Percent Similarity: 88.62% Conserved: 27
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 40.23% Indels: 1
Gaps: 1
US-10-029-345a-108_COPY_538_1443 (1-906) x US-09-619-380-16 (1-170)

QY 400 GGGCTCTGTGAAGAAAA--TCACCTTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTA 456
DB 1 G1yleuCyglucllylsProAlaAlaLeuLeuProMeSerleuSerInProCySleu 20
QY 457 CCTGTGGCAACATGGGGCCAAACCGCAATTTCTCCCAATTTTATCTTGCTGCCAGCA 516
DB 21 ProValProSerValG1yleuThrArgIleLeuProH1leuYrleuGlySerGlnlys 40
QY 517 GATGTCTCAACAAGAGCTGATATACAGCAAGATGGATGGTATGTGTTAAATGCCAGC 576

```
Db 41 AspvallleuAsnlysaPleuMeetmrgInaSnGlyIleSerTyValleuAsnLaser 60
Qy 577 TATACCTGTCAGAAAGCTGACTTATATCCCGAGTCTCATTCCTCCGCTGCTGTAAT 636
Db 61 AsnserCyseProlysaProaPheIleCySgluSerArgPheMetValaProIleAsn 80
Qy 637 GACAGCTTTTGTAGAAATTTTGGCGGTGGCAATCAATGATGATTTTCAATGAGAAA 696
Db 81 AspnsmTyrcySgluIlyleuSerProIleuAsnlysaPheIleGluPheIleAsnly 100
Qy 697 GCAAAAGCTCCCAATGATGATGTTCTAGTGCATCTTTAGCTGGAGATCTCCGCTCCGC 756
Db 101 AlalylyseuSerCyseInaIleValaHscYseuIleaglIleSerArgSerIle 120
Qy 757 ACCATGCTATGCTGCTACATCATGAGAGATGAGCATGCTTAAATGATGATGAGAGA 816
Db 121 ThrIleAlaIleAlaIleAlaIleIleMetIleThreMetIleMetSerArgPheIle 140
Qy 817 TTTGTGAAAGAAAAGACCTACTATATCTCCAACTTCAATTTTCTGGGCAACTCTG 876
Db 141 PheValIysaPheArgProSerIleSerProaPheAsnPheIleuclIyInleu 160
Qy 877 GACATGAGAGAAAGATTAG 897
Db 161 GluTyrgluArgThrIleuIys 167
```

RESULT 6

```
US-09-557-921-2
/ Sequence 2, Application US/09557921
/ Patent No. 6551810
/ GENERAL INFORMATION:
/ APPLICANT: Luche, Ralf M.
/ TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
/ FILE REFERENCE: 200125.416
/ CURRENT APPLICATION NUMBER: US/09/557,921
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 482
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-557-921-2
```

Alignment Scores:

```
Pred. No.: 2,15e-48
Score: 469.00 Length: 482
Percent Similarity: 57.72% Matches: 107
Best Local Similarity: 35.91% Conservative: 65
Query Match: 28.83% Mismatches: 96
Indels: 30
Gaps: 7
```

US-10-029-345a-108_COPY_538_1443 (1-906) x US-09-557-921-2 (1-482)

```
Qy 79 CTGCTAATTGATAGCGGCAATTTGTGAATACATATACATCCCAATTTTGGAGCCAT 138
Db 173 ValIleIleAspCyArgProPheMetGluTyAsnlySerHisIleGlnIleVala 192
Qy 139 AATATCACTCTCTC--AAGCTTATGAGCGAAGGTGGACAGACAAAGCTTAAAT 195
Db 193 HisIleAsnCySalasPlyIleSerArgArgIleGlnIleGlnIlyValIleThrVal 212
Qy 196 ACAAGAGCTCATCGA--CATTCAGCGAAACATAGGTTGACATTGATGATGACAGAG 252
Db 213 LeuAspIleuIleSerCyArgGluIlyAspSerPheIysArgIlePheSerIyGlu 232
Qy 253 GTTGTACTTACATGAAAGCTCCCAAGAGTGTGCTCTCTCTTCAAGACTTTTCTC 312
Db 233 IleIleValIlyAspGluAsnThrAsnGluProSerArgValMetProSerGlnProIle 252
Qy 313 ACTGTACTTGTGGTAACTGAGAAAGACTTCAACTGTGTTCACTGCTTGGAGGTGGG 372
```

```
Db 253 HisIleValleuIleuSerIleuIyAspGluIlyArgIleuProIleuValleuIySglY 272
Qy 373 TTTGCTAGATTTCTCTGTTGTTTCCCTGCGCTCTGTGAAGAGAAA----- 417
Db 273 LeuSerSerPheIySgluAsnHisIleGluAsnIyCysAspAsnSerIleuIleuGlnI 292
Qy 418 -----TCACCTCTAGCCCTACCTGATTTCT 444
Db 293 CysArgGluValaIleGlyIleValaSerAlaIleSerIleuProIleuProIle 311
Qy 445 CAGCTTGGCTTACCT--GTTGCCACATTTGGGCAACCCGAATTTTCCCATCTTAT 501
Db 312 ---ProIleThrProaPheIleGluAsnAlaIleuIleuThrProIleuProIleuPhe 330
Qy 502 CTGGCTCCGCGAGAGATGCTCCACAGAGACTGATACAGACAGATGGAGTTGTTAT 561
Db 331 leuGluIyAsnGluIyAsnAlaIleuAspIleuAspThrMetGlnArgIleuAsnIleIy 350
Qy 562 GTTGTAAATGCCAC-----TATACCTGTCCAAAGCTGACTTATCCCC 606
Db 351 ValIleAsnValIleThrThrIleuProIleuTyHisIyrgIlySglYleuPhe----- 368
Qy 607 GAGTCTCATTTCTCGGTGGTGGCTGTGAATGACAGCTTTTGTGAGAAAATTTGCCGTG 666
Db 369 -----AsnTyrcIyAsnArgIleuProIleThrAspSerAsnlySgluAsnIleuArgIleu 386
Qy 667 TTGACAAATCAAGTAAATTTTCAATTTGAGAAAGCAAAAGCTCCCAATGATGATGTTAGT 726
Db 387 PheGlnIleuAlaPheGluPheIleGluIleuAlaHisGlnIySglYleuIle 406
Qy 727 CACTGTTTAACTGGAGATCTCCGCTCCGCGCAACATGCTATGCTTACATCAAGAAG 786
Db 407 HisCySgluIleaglIyAlSerArgSerIleThrIleValIleAlaIyIleuMetIyHis 426
Qy 787 ATGACATGCTTGTAGATGAGCTTACAGATTTGTGAAAGAAAAGCACTTATATCT 846
Db 427 ThrArgMetIleuMetIleuAspAlaIyIySphValIySglYleuArgProIleIleSer 446
Qy 847 CCAACTTCAATTTTCTGGGCAACTCTGAGCTAGTAAAGAGATTAAAGAC 900
Db 447 ProAsnIleuAsnPheMetGlyIleuIleuGluIleuGluIleuAsn 464
```

RESULT 7

```
US-08-530-290-23
/ Sequence 23, Application US/08530290
/ Patent No. 5958721
/ GENERAL INFORMATION:
/ APPLICANT: Marshall, Christopher John
/ APPLICANT: Ashworth, Alan
/ TITLE OF INVENTION: Methods for Screening of Substances for
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ FILING DATE: 14-DEC-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO PCT/GB94/00694
/ FILING DATE: 31-MAR-1994
```

```
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9402573.1
FILING DATE: 10-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9307250.2
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Baselian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 084611-000000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-530-290-23

Alignment Scores:
Pred. No.: 2,73e-43 Length: 394
Score: 427.00 Matches: 101
Percent Similarity: 52.88% Conservative: 64
Best Local Similarity: 32.37% Mismatches: 116
Query Match: 26.24% Indels: 31
DB: Gaps: 8

US-10-029-345a-108_COPY_538_1443 (1-906) x US-08-530-290-23 (1-394)
QY 45 GGTTGCTCTGCTGGAAGTGAACGGA-----AAAGT 77
DB 26 Gtlyglalaglyly-SerGlyserhlsGlythrlleuglyleuproserylygllyscy 45
QY 78 GCTGCTAATGATGAGCCGCAATTTGGAATCAATACATCCCACTTTTGGAGCCAT 137
DB 45 elenleuLeuAapCyArGPrroPheulaahlsSerAlaglytyrilleuGlyserVa 65
QY 138 TAATATCAACTGCTCCAGCTTATGAGCGAGGTTGCAACAGACAAAGTTAAATTAC 197
DB 65 lAnValArgCyAsnThrlleValArgArg--AlaGlySerValserleuG 84
QY 198 AGAGCTATC-----CAGCATTCAGCGAAACATAAGTTGACATTGATTCAGTCAGAA 251
DB 84 uGlnlleleuProAlaglnluGlnValArgAlaArgleuArgSerGlyleuTyrsAl 104
QY 252 GGTGTAGTTAGATCAAGCTCCCAAGATGTTGCTCTCTCTTCAAGCTGTTTCT 311
DB 104 aValilleValtyrAspGlnArgSerProArgAlaGlnSerleuArgGlnuAspserThryA 124
QY 312 CACTGTACTTCTGGTAACTGGAAGAGCTTC-----AACTGTGTTCACTGCTGC 365
DB 124 lSerleuValValGlnAlaLeuArgArgAenAlaGlnArgThrAspIleCySerleu 144
QY 366 AGGTGGTTTGCAGATTCTCTGTTGTTTCCCTGGCTCTGGAAGAAATCCACTCT 425
DB 144 sGlygllyTyGlnArgPheSerSerGlnTyPrroGlnPheCySerlyerThryAlaLe 164
QY 426 A-----GTCCCTACCTGCAATTTCTCAGCCCT-----TGCTT 455
DB 164 uAlaAlaIleProProProValAlProProSerAlaThGlnProleuAspLeuAspCySe 184
QY 456 ACCTGTGTC-----AACATTGGGCCCAACCCGAATTTCTTCCCATCTTGA 500
DB 184 rSerCyGlyThrProleuhtsAspGlnGlnuGlyProValGlnIleleuProheleuTy 204
QY 501 TCTTGGCTGCAGAGATGCTCTCAACAAGAGCTGATCAGACAGATGGATGGTTA 560
DB 204 rleuGlySerAlaTythlsAlaAlaArgAspMetleuAspAlaLeuGlyleuThral 224
QY 561 TGTTTAAATGACAGTATACCTGTCCAAAGCCTGACTTATCCCGAGTCAATTTCT 620
```

```
Db 224 aleuLeuAenValSerSerAspCyProAsn---HisPheGlnGlyHisTyrlnTyry 243
QY 621 GCCTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTCCGTGTGGACAAATCACT 680
DB 243 sCyrlleProValGlnuAspAsnhtlsAlaAspIleSerSerTrpPheMetGlnuAl 263
QY 681 AGATTTCATTGGAAGCAAAAGCCCTCAATGATGTTGTTCTGACACTTTAGCTGG 740
DB 263 eGlnTyrlleAspAlaVallyAspCyArGlyArgValleuValhtlscyGlnuAl 283
QY 741 GATCTCCGCTCCGACCATCGCTATCGCTTACATCATGAGAGATGACATGCTTT 800
DB 283 ylleSerArgSerAlaThrlleCySerleuAlTyrlleuMetMetlyleValArgle 303
QY 801 AGATGAAGCTTACAGATTTGTGAAGAAAAGAACTACTATATCTCAACTTCAATT 860
DB 303 uGlnGlnAlaPheGlnPheVallySerGlnArgSerllelleSerProAsnPheserPh 323
QY 861 TCTGGGCCAATCTGAGCATATGAGAAAGATT 894
DB 323 ewetGlyGlnleuGlnPheGlnuSerGlnVal 334

RESULT 8
US-09-702-705-805
; Sequence 805, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodges, Michael A.
; APPLICANT: Ranger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retler, Marc
; APPLICANT: Mannon, Jane
; APPLICANT: Pan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-705-805

Alignment Scores:
Pred. No.: 2,73e-43 Length: 394
Score: 427.00 Matches: 101
Percent Similarity: 52.88% Conservative: 64
Best Local Similarity: 32.37% Mismatches: 116
Query Match: 26.24% Indels: 31
DB: Gaps: 8

US-10-029-345a-108_COPY_538_1443 (1-906) x US-09-702-705-805 (1-394)
QY 45 GGTTGCTCTGCTGGAAGTGAACGGA-----AAAGT 77
DB 26 Gtlyglalaglyly-SerGlyserhlsGlythrlleuglyleuproserylygllyscy 45
QY 78 GCTGCTAATGATGAGCCGCAATTTGGAATCAATACATCCCACTTTTGGAGCCAT 137
DB 45 elenleuLeuAapCyArGPrroPheulaahlsSerAlaglytyrilleuGlyserVa 65
QY 138 TAATATCAACTGCTCCAGCTTATGAGCGAGGTTGCAACAGACAAAGTTAAATTAC 197
DB 65 lAnValArgCyAsnThrlleValArgArg--AlaGlySerValserleuG 84
QY 198 AGAGCTATC-----CAGCATTCAGCGAAACATAAGTTGACATTGATTCAGTCAGAA 251
```

```

Db      84 uGnIleuProAlaGluGluValAlaArgAlaArgLeuSerGlyLeuTyrSerAl 104
QY      252 GGTGTGATTTAGCAACAAAGCTCCAGAGATGTTGCTCTCTCTCTTCCAGCTGTTTCT 311
Db      104 aValIleValIlyrAspAlaArgSerProAlaGluSerLeuArgIleuAspSerThrVa 124
QY      312 CACTGACTTCTGGGTAACTGGAGAGAGCTTC-----AAGCTGTCACTGCTTGC 365
Db      124 IserLeuValIAlaGlnAlaLeuArgAsnAlaGlnArgThrAspIleCysIleuLeu 144
QY      366 AGGTGGTTTGTGAGTCTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
Db      144 sGlyGlyTyrGlnArgPheSerSerGlyTyrProGluPheCysSerIleThrAlaLe 164
QY      426 A-----GTCCCTACCTGATTTCTCGAGCT-----TGCTT 455
Db      164 uAlaAlaIleProProProValProProSerAlaThrGluProLeuAspLeuAspCysSe 184
QY      456 ACCTGTTGCC-----AACATTGGGCCAACCCGAATTCTTCCCAATCTTTA 500
Db      184 rSerCysGlyThrProLeuHisAspGlnGluValArgAlaArgArgAspMetLeuAspAla 204
QY      501 TCTTGAGCTGCACGAGATGCTCTCAACAGAGAGCTGATACAGAGATGGAGATTGGTTA 560
Db      204 rLeuGlySerAlaTyrHisAlaAlaArgArgAspMetLeuAspAlaLeuGlyIleThrAl 224
QY      561 TGTGTTAAATGCCAGCTATACCTGCAAGAGCTGATTTATCCCGAGTCTGATTTCCCT 620
Db      224 AlauLeuAsnValIserSerAspCysProAsn--HisPheGlnGlyHisTyrGlnTyrI 243
QY      621 GCGTGGCTGTGAATGACAGCTTTTGTGAGAAATTTTCCCGTGTGAGCAAAATCACT 680
Db      243 sCysIleProValGlnAspAsnHisIlyAlaAspIleSerSerIlePheMetGlnAlaI 263
QY      681 AGATTTCAATTAGAAAGCAAAAGCTCAATGATGTTGTTGCTGCTGCTGCTGCTGCTG 740
Db      263 eGlnTyrIleAspAlaValIlyAspCysAspGlyArgValLeuValHisCysGlnAla 283
QY      741 GATTCCTCCGCTCCGCAACATCGCTTATCGCTTACATGATGAGAGATGAGATGATGCT 800
Db      283 yIleSerArgSerAlaThrIleCysLeuAlaTyrLeuMetCetylAspValArgLe 303
QY      801 AGATGAAGCTTACAGATTTGTGAAAGAAAAAGACCTACTATATCTCCAAATCTCAATT 860
Db      303 uGlnGlnAlaPheGluPheValIlySerGlnArgArgSerIleIleSerProAsnPheSer 323
QY      861 TCTGGGCACTCTCGAGCTATGAGAGAAATTT 894
Db      323 eMetGlyGlnLeuLeuGlnPheGluSerGlnVal 334

RESULT 9
US-09-702-705-827
/ Sequence 827, Application US/09702705
/ Patent No. 6504010
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongrong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Panger, Gary
/ APPLICANT: Vedrick, Tom
/ APPLICANT: Carter, Darick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C14
/ CURRENT APPLICATION NUMBER: US/09/702, 705
/ NUMBER OF SEQ ID NOS: 2000-10-30
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 827

```

```

/ LENGTH: 394
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-09-702-705-827

Alignment Scores:
Pred. No.: 2,73e-43
Score: 427.00
Percent Similarity: 53.85%
Best Local Similarity: 32.69%
Query Match: 26.24%
DB: 4
Gaps: 9

US-10-029-345a-108_COPY_538_1443 (1-906) x US-09-702-705-827 (1-394)
QY      45 GGTGGCTGTGCTGCAAGATGGAACGGA-----AAAGT 77
Db      26 GlyIlyAlaGlyly-SerGlySerHisGlyIlyLeuProSerGlyGlyIlyCys 45
QY      78 GCTCTAATGATGATGCGGCAATTTGTGAAATACATACATCCCATTTTGAAGCAT 137
Db      45 sLeuLeuAspCysArgProPheLeuAlaHisSerAlaGlyIlyrIleLeuGlySerVa 65
QY      138 TAATATCACTGCTCCAGCTTATGAGAGAGGTGCAACAGAGCAAAAGTTAATTAC 197
Db      65 laenValArgCysAsnThrIleValArgArg--AlaIlySerValIserLeuG 84
QY      198 AGAGCTCATC-----CAGCATTCAGCGAAACATTAAGTTGACATTTGACAGTCAGAA 251
Db      84 uGlnIleuProAlaGluGlnGluValAlaArgAlaArgLeuArgSerGlyLeuTyrSerAl 104
QY      252 GGTGTGATTTAGCAACAAAGCTCCCAAGATGTTGCTCTCTCTCTTCCAGCTGTTTCT 311
Db      104 aValIleValIlyrAspAlaArgSerProAlaGluSerLeuArgIleuAspSerThrVa 124
QY      312 CACTGACTTCTGGGTAACTGGAGAGAGCTTC-----AAGCTGTCACTGCTTGC 365
Db      124 IserLeuValIAlaGlnAlaLeuArgArgAsnAlaGlnArgThrAspIleCysIleuLeu 144
QY      366 AGGTGGTTTGTGAGTCTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
Db      144 sGlyGlyTyrGlnArgPheSerSerGlyTyrProGluPheCysSerIleThrAlaLe 164
QY      426 A-----GTCCCTACCTGATTTCTCGAGCT-----TGCTT 455
Db      164 uAlaAlaIleProProProValProProSerAlaThrGluProLeuAspLeuAspCysSe 184
QY      451 ----TGC---TTACCTGTTGCCAATTT---GGGCAACCCGAATTTCTTCCCAATCTTTA 500
Db      184 rSerCysGlyThrProLeuHisAspGlnGluValArgAlaGlnIleuProPheLeu 204
QY      501 TCTTGAGCTGCACGAGATGCTCTCAACAGAGAGCTGATACAGAGATGGAGATTGGTTA 560
Db      204 rLeuGlySerAlaTyrHisAlaAlaArgArgAspMetLeuAspAlaLeuGlyIlyIleThrAl 224
QY      561 TGTGTTAAATGCCAGCTATACCTGTCMAAGCCTGATTTATCCCGAGTCTGATTTCCCT 620
Db      224 AlauLeuAsnValIserSerAspCysProAsn--HisPheGlnGlyHisTyrGlnTyrI 243
QY      621 GCGTGGCTGTGAATGACAGCTTTTGTGAGAAATTTTCCCGTGTGAGCAAAATCACT 680
Db      243 sCysIleProValGlnAspAsnHisIlyAlaAspIleSerSerIlePheMetGlnAlaI 263
QY      681 AGATTTCAATTAGAAAGCAAAAGCTCCCAATGATGTTTGTGAGCAATTTTGAAGCAT 740
Db      263 eGlnTyrIleAspAlaValIlyAspCysAspGlyArgValLeuValHisCysGlnAla 283
QY      741 GATTCCTCCGCTCCGCAACATCGCTATGCGCTTACATGATGAGAGATGAGATGCTT 800
Db      283 yIleSerArgSerAlaThrIleCysLeuAlaTyrLeuMetCetylAspValArgLe 303
QY      801 AGATGAAGCTTACAGATTTGTGAAAGAAAAAGACCTACTATATCTCCAAATCTCAATT 860

```

Db 303 ugiugluAlaPhegluPheVallyseGlnArgSerlelleSerProAnpBheserPh 323
QY 861 TCTGGCCCACTCTGCATATAGAAAGATT 894
Db 323 ewetGlyInLeuIngleInpHeGInserGlnVal 334

RESULT 10
US-09-736-457-805
Sequence 805, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodges, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Pan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FaSTSeq for Windows Version 3.0
SEQ ID NO 805
LENGTH: 394
TYPE: PRN
ORGANISM: Homo sapiens
US-09-736-457-805

Alignment Scores:
Pred. No.: 2,73e-43 Length: 394
Score: 427.00 Matches: 101
Percent Similarity: 52.88% Conservative: 64
Best Local Similarity: 32.37% Mismatches: 116
Query Match: 26.24% Indels: 31
Gaps: 8
DB: 4

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-736-457-805 (1-394)

QY 45 GGTCGCTCTCTCGAAGTGAAGCGAA-----AAAGT 77
Db 26 G1yG1yA1aG1yG1y-SerG1ySerH1seG1yHrLeuG1yLeuProSerG1yG1ySeCy 45
QY 78 GGTGCTAATGATGAGCGCCATTTGGGAATACATATACATCCACATTTTGGAGCGAT 137
Db 45 aleuLeuLeuAspCyArGrProPheLeuAlaH1seAlaG1yTyrH1leuG1ySeVa 65
QY 138 TATATCAACTGCTCCAGCTTGAAGCGAGTTGCAACAGAGAAAGTTTAATTAC 197
Db 65 laenValaVgCyAenThrH1leValaArgaGrg---AlaYseG1ySerAlseValseuG1 84
QY 198 AGAGCTCATC-----CAGCATTCAGCGAAACATAAGCTTGACATTGTCAGTGA 251
Db 84 uGlnH1leuProAlaGlnGlnGlnValaArgAlaArgLeuArgSerG1yLeuTyrSeAl 104
QY 252 GGTGAGATTAGCATCAAGCTCCCAAGATGTGCTCTCTTCAAGACTGTTTTCG 311
Db 104 aValH1leValaTyrAspGlnArgSerProAlaGlnUserLeuArgGlnAspSerThrVa 124
QY 312 CACTGACTTCTGGTAAACTGAGAAAGACTTC-----AACTGTTCACTGCTTGC 365
Db 124 lserleuValaValaGlnAlaLeuArgaRenaAlaGlnArgThrAspH1leCySeLeuLeu 144
QY 366 AGGTGGTTTGCAGAGTCTCTGTTGTTTCCCTGGCGCTCTGGAAGAAATCCACTCT 425
Db 144 sG1yG1yTyrGlnArgPheSerSerGlnTyrProGlnPheCySeSerH1leThrYsaAla 164
QY 426 A-----GTCCCTACCTGATTTCTCAGCCT-----TGCTT 455

Db 164 uAlaAlaH1leProProProValProProSerAlaThrGlnProLeuAspLeuAspCySe 184
QY 456 ACCTGTGCC-----AACTGGGCCCAACCCGAATTTCCCAATCTTTA 500
Db 184 rSerCyG1yHrProLeuH1seArgInGlnGlnProValaGlnH1leuProPheLeu 204
QY 501 TCTTGGCTGCAGCGAGATGTCTCAACAGAGCTGATACAGCAGATGGATGGTTA 560
Db 204 rLeuG1ySerAlaTyrH1aAlaAlaArgaRgAspMetLeuAspAlaLeuG1yH1eHrAl 224
QY 561 TGTGTTAAATGCCAGTATATACCTGTCMAAGCTGATTAATCCGAGCTCATTTCT 620
Db 224 aleuLeuAenValaSerSerAspCyProAn---H1sePheGlnG1yH1eTyrGlnTyr 243
QY 621 GCCTGCTCTGTCATGACAGCTTTTGTGAGAAATTTTGCCTGGTGGACAAATCACT 660
Db 243 vCyH1leProValaGlnAspAsnH1leYsaAlaAspH1leSerSerTrpPheMetGlnAla 263
QY 681 AGATTTCATTGAGAAAGCAAAAGCCTCAATGATGATGTTCTTATGACATGTTAGCTGG 740
Db 263 eGlnTyrH1leAspAlaValaH1seAspCyArgGlnArgValaH1eCySeGlnAlaG1 283
QY 741 GATTCCTCCGCTCGCCACCATGCTATCGCTATCATATGAAAGAGATGACATGCTTT 800
Db 283 yH1seSerArgSerAlaThrH1leCySeuAlaTyrH1eUmetMetLeuYsaGlnArgLe 303
QY 801 AGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTACTATATCTTCAACTTCATT 860
Db 303 ugiugluAlaPhegluPheVallyseGlnArgSerlelleSerProAnpBheserPh 323

QY 861 TCTGGCCCACTCTGCATATAGAAAGATT 894
Db 323 ewetGlyInLeuIngleInpHeGInserGlnVal 334

RESULT 11
US-09-736-457-827
Sequence 827, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodges, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Pan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FaSTSeq for Windows Version 3.0
SEQ ID NO 827
LENGTH: 394
TYPE: PRN
ORGANISM: Homo sapiens
US-09-736-457-827

Alignment Scores:
Pred. No.: 2,73e-43 Length: 394
Score: 427.00 Matches: 102
Percent Similarity: 53.85% Conservative: 66
Best Local Similarity: 32.69% Mismatches: 113
Query Match: 26.24% Indels: 31
Gaps: 9
DB: 4

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-736-457-827 (1-394)


```
Db 243 bcysileProvalGlubspashhlsylsalaeserSerrtrPheMeGlual1 263
Qy 681 AGATTTCATTGAGAAAGCAAAAGCCTTCAATGATGTGTTCTGTCACCTGTTTACTGG 740
Db 263 egutyrileabpAlavalysabpCyabrglyarGvalleuvalHieCyseGlmlaGl 283
Qy 741 GATTCCTCCGCTCCGACCATGCTATCGCTCATCATCAAGAGATGACATGCTTT 800
Db 283 ylleSerhargseralathrlleCyseuvalatryleuMeMeLylabrgvalargyle 303
Qy 801 AGATGAGCTTACAGATTGTGTGAGAAAGAAAGCCTACTATATCTCAAACTTCATT 860
Db 303 ugluglualphegluPhegluPhegluInrargserlleleSerProhsheserPh 323
Qy 861 TCTGGCCCAACTCTGACATATAGAAAGATT 894
Db 323 ewetGlyInleuGlInPhegluSerGlInVal 334

RESULT 13
US-09-614-124B-827
/ Sequence 827, Application US/09614124B
/ Patent No. 6630574
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mammon, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ FILE REFERENCE: 210121.478C9
/ CURRENT APPLICATION NUMBER: US/09/614,124B
/ CURRENT FILING DATE: 2001-07-11
/ NUMBER OF SEQ ID NOS: 1668
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 827
/ LENGTH: 394
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-614-124B-827

Alignment Scores:
Pred. No.: 2,73e-43 Length: 394
Score: 427.00 Matches: 102
Percent Similarity: 53.85% Conservative: 66
Best Local Similarity: 32.69% Mismatches: 113
Query Match: 26.24% Indels: 31
DB: 4 Gaps: 9

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-614-124B-827 (1-394)
Qy 45 GGTGGCTGCTGGAGAGGAGAGGAA-----AAAGT 77
Db 26 GlyelYAlGlyly-SerilyserihseGlythrlleuGlyleuProserGlyGlylyscY 45
Qy 78 GGTGCTAATGTGATAGCGGCGCATTTGTGATATACATATCCACATTTTGGAGCCAT 137
Db 45 aleuLeuLeuabpCyabrgProheleuLahlseserlaglYtrYlleuGllyserVa 65
Qy 138 TAATATCACTGCTCCAGCTTATGAAAGGAGGTGCAACAGCAAGAAAGTAAATTC 197
Db 65 lhenValargCyabshthrlleValargarg--AlalyGlyserValserleuGl 84
Qy 198 AGAGCTCATC-----CAGCATTCAGGAAACATTAAGTTCATTCATTCAGTCAGAA 251
Db 84 ugnlileleuProAlgluGlulgluValargAlahrgleuargserGlyleuYrserAl 104
Qy 252 GGTGTAGATTACATCAAGACTCCCAAGATGTGGCTCTCTCTTCAGACATGTTTCT 311
Db 104 aValileVallyrAbpGlubrgserProargAlagluserleuAbrglubsperthrya 124

US-09-671-325-805
/ Sequence 805, Application US/09671325
/ Patent No. 6667154
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mammon, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C12
/ CURRENT APPLICATION NUMBER: US/09/671,325
/ CURRENT FILING DATE: 2000-09-26
/ NUMBER OF SEQ ID NOS: 1825
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 805
/ LENGTH: 394
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-671-325-805

Alignment Scores:
Pred. No.: 2,73e-43 Length: 394
Score: 427.00 Matches: 102
Percent Similarity: 53.85% Conservative: 66
Best Local Similarity: 32.69% Mismatches: 113
Query Match: 26.24% Indels: 31
DB: 4 Gaps: 9

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-614-124B-827 (1-394)
Qy 312 CACTGACTCTTGGGTAACGTGAGAGAGCTTC-----AACTGTTCACTGCTGTC 365
Db 124 lserleuValValGlmlaAleuargabshhlaGlualgthrlabrpilleCyseleuY 144
Qy 366 AGGTGGTTTGGCTGAGATTCTCTGTTGTTCCCTGGCTCTGTGAGAGAAATCCACTCT 425
Db 144 eGlYglYrGluabrgPheSerSerleuYrProgluPheCyseYserlyethryalale 164
Qy 426 A-----GTCCCTACTGTCATTTCTCAGCCT----- 450
Db 164 uAlaAlaleProProProvalProProseralathrgluProleuabpLeuGlyCyse 184
Qy 451 -----TTCCTGTTGCCAACTT---GGGCCAACCCGAATTCCTCCCAATCTTA 500
Db 184 rserYsglythrProleuHlsabpInglYglProvalGlulileuLeuPheleuY 204
Qy 501 TCTTGCTGTCAGCGAGATGTCCTCAACAGAGCTGATACAGCAATGAGATTTGGTTA 560
Db 204 rleuGlYserAlatYrhlalAlalargharghmetleuabpAlaleuGlYlethrl 224
Qy 561 TGTGTTAAATGCGAGCTATACCTGTCCAAAGCCTGATTTATCCCGAGTCTCATTTCT 620
Db 224 aleuLeuabpValserSerabpCyseProsh--HisPhegluGlYhlleYrGlInYrly 243
Qy 621 GCGTGGCCTGTGAATGACAGCTTTGTGAGAAATTTTCCGCTGTTGAGCAAAATCAGT 680
Db 243 bcysileProvalGlubspashhlsylsalaeserSerrtrPheMeGlual1 263
Qy 681 AGATTTCATTGAGAAAGCAAAAGCCTTCAATGATGTGTTCTGTCACCTGTTTACTGG 740
Db 263 egutyrileabpAlavalysabpCyabrglyarGvalleuvalHieCyseGlmlaGl 283
Qy 741 GATTCCTCCGCTCCGACCATGCTATCGCTCATCATCAAGAGATGACATGCTTT 800
Db 283 ylleSerhargseralathrlleCyseuvalatryleuMeMeLylabrgvalargyle 303
Qy 801 AGATGAGCTTACAGATTGTGTGAGAAAGAAAGCCTACTATATCTCAAACTTCATT 860
Db 303 ugluglualphegluPhegluPhegluInrargserlleleSerProhsheserPh 323
Qy 861 TCTGGCCCAACTCTGACATATAGAAAGATT 894
Db 323 ewetGlyInleuGlInPhegluSerGlInVal 334

RESULT 14
US-09-671-325-805
/ Sequence 805, Application US/09671325
/ Patent No. 6667154
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mammon, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C12
/ CURRENT APPLICATION NUMBER: US/09/671,325
/ CURRENT FILING DATE: 2000-09-26
/ NUMBER OF SEQ ID NOS: 1825
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 805
/ LENGTH: 394
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-671-325-805

Alignment Scores:
Pred. No.: 2,73e-43 Length: 394
Score: 427.00 Matches: 102
Percent Similarity: 53.85% Conservative: 66
Best Local Similarity: 32.69% Mismatches: 113
Query Match: 26.24% Indels: 31
DB: 4 Gaps: 9
```

Pred. No.: 2,73e-43
 Score: 427.00
 Percent Similarity: 52.88%
 Best Local Similarity: 32.37%
 Query Match: 26.24%
 DB: 4
 Matches: 394
 Conserved: 101
 Mismatches: 64
 Indels: 116
 Gaps: 8

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-671-325-805 (1-394)

```

45 GGTGCTGCTGCTGGAAGTGAACGGA-----AAAGT 77
26 GlyGlylaaglyly-SerGlySerHisgLyThrLeuGlyLeuProSerGlyGlyLysCy 45
78 GCTGCTAATGATAGCCGCGCATTTTGGAATCAATACATCCCATTTTGAAGCAT 137
45 sLeuLeuLeuAspCysArgProPheLeuLahisSerlaaglyTyrlleuGlySerVa 65
138 TAATATCACTGCTCAAGCTTATGAAAGAGCTTGCAACAGACAAAGTTAATTAC 197
65 laenValArgCysAsnThrileValArgArgArg---AlaYsglySerValSerLeuG 84
198 AGAGCTCANTC-----CAGCATTCAGCGAACAATAGAGTTGACATTGACGTCA 251
84 uGlnIleuProAlaGluGluGluValArgAlaArgLeuArgSerGlyLeuTySerAl 104
252 GGTGTAATGATTAAGATCAAAAGCTCCCAAGATGTTGCTCTCTTCAAGATGTTTCT 311
104 aValIleValIYrAspGluArgSerProArgAlaGluSerLeuArgGluAspSerTh 124
312 CACTGACTCTGCGGTAACTGGAAGAGCTTC-----AACTGTTCACTGCTTCC 365
124 sLeuLeuValAlGlnAlaLeuArgArgAsnAlaGluArgThAspIleCysLeuLeu 144
366 AGTGGGTTTGCTGAGTTCTCTCGTTGTTCCCGGCTCTGGAAGAAATCCACTCT 425
144 sGlyGlyTyrgluArgPheSerSerGlyTyrgluPheCysSerIysThrIleAla 164
426 A-----GTCCCTACCTGCATTTCTCAGCC-----TGCTT 455
164 uAlaAlaIleProProProValProProSerAlaThrGluProLeuAspLeuAspC 184
456 ACCTGTTGCC-----AACATGGGCGCAACCCGAATTTCTTCCCAATCTTTA 500
184 rSerCysGlyThrProLeuHisAspGlnGluGlyProValGluIleuProPheLeu 204
501 TCTTGCTGTCAGCAGATGCTCTCAACAGAGCTGATACAGCAGATGGATTTGTTA 560
204 rLeuGlySerAlaTyrlleAlaAlaArgArgAspMetLeuAspAlaLeuGlyIle 224
561 TGCTTAAATGCCAGCTATACCTGTCGAAGCTGACTTTATCCCGAGTCTCATTTCT 620
224 aLeuLeuAsnValSerSerAspCysProAsn---HisPheGluGlyHisTyrglu 243
621 GCGTGTGCTGGAATGAACGCTTTTGGAATAATTTGCGGTGTCGACAATCAGT 680
243 sCysIleProValGluAspAsnHisIysAlaAspIleSerSerThrPheMetGluAla 263
681 AATATTCAATGAAAGAAAGAAAGCTCCCAATGATGTTGTTCTAGTCACTGTTTAC 740
263 eGluTyrlleAspAlaValIysAspCysArgGlyArgValLeuValHisCysGlu 283
741 GATCTCCGCTCGCGCAACATCGCTATCCGCTTCAATCAATGAAGATGATGATCTTT 800
283 yIleSerArgSerAlaThrIleCysLeuAlaTyrlleuMetMetIysLysArgVal 303
801 AGATGAAGCTTACAGATTTTGAAAGAAAGAAAGCTTACTATATCTCCAAATTTA 860
303 uGluGluAlaPheGluPheValIysGluArgIleGlySerIleIleSerProAsn 323
861 TCTGGGCAACTCCTGAGACTATGAGAAAGATTT 894
323 eMetGlyGlnLeuLeuGlnPheGlnSerGlnVal 334

```

RESULT 15
 US-09-671-325-827
 / Sequence 827, Application US/09671325
 / Patent No. 6667154
 / GENERAL INFORMATION:

```

/ APPLICANT: Wang, Tongrong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Pan, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C12
/ CURRENT APPLICATION NUMBER: US/09/671,325
/ NUMBER OF SEQ ID NOS: 1825
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 827
/ LENGTH: 394
/ TYPE: PRT
/ ORGANISM: Homo sapiens

```

US-09-671-325-827

Alignment Scores:

Pred. No.: 2,73e-43
 Score: 427.00
 Percent Similarity: 53.85%
 Best Local Similarity: 32.69%
 Query Match: 26.24%
 DB: 4
 Matches: 394
 Conserved: 102
 Mismatches: 66
 Indels: 113
 Gaps: 9

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-671-325-827 (1-394)

```

45 GGTGCTGCTGCTGGAAGTGAACGGA-----AAAGT 77
26 GlyGlylaaglyly-SerGlySerHisgLyThrLeuGlyLeuProSerGlyGlyLysCy 45
78 GCTGCTAATGATAGCCGCGCATTTTGGAATCAATACATCCCATTTTGAAGCAT 137
45 sLeuLeuLeuAspCysArgProPheLeuLahisSerlaaglyTyrlleuGlySerVa 65
138 TAATATCACTGCTCAAGCTTATGAAAGAGCTTGCAACAGACAAAGTTAATTAC 197
65 laenValArgCysAsnThrileValArgArgArg---AlaYsglySerValSerLeuG 84
198 AGAGCTCANTC-----CAGCATTCAGCGAACAATAGAGTTGACATTGACGTCA 251
84 uGlnIleuProAlaGluGluGluValArgAlaArgLeuArgSerGlyLeuTySerAl 104
252 GGTGTAATGATTAAGATCAAAAGCTCCCAAGATGTTGCTCTCTTCAAGATGTTTCT 311
104 aValIleValIYrAspGluArgSerProArgAlaGluSerLeuArgGluAspSerTh 124
312 CACTGACTCTGCGGTAACTGGAAGAGCTTC-----AACTGTTCACTGCTTCC 365
124 sLeuLeuValAlGlnAlaLeuArgArgAsnAlaGluArgThAspIleCysLeuLeu 144
366 AGTGGGTTTGCTGAGTTCTCTCGTTGTTCCCGCTCTGGAAGAAATCCACTCT 425
144 sGlyGlyTyrgluArgPheSerSerGlyTyrgluPheCysSerIysThrIleAla 164
426 A-----GTCCCTACCTGCATTTCTCAGCC-----TGCTT 450
164 uAlaAlaIleProProProValProProSerAlaThrGluProLeuAspLeuGlyC 184
451 ----TGCTTACCTGTCGAACAT---GGGCAACCCGAATTTCTTCCCAATCTTTA 500
184 rSerCysGlyThrProLeuHisAspGlnGlyGlyProValGluIleuProPheLeu 204

```

```
QY 501 TCTTGCTGCACGACGATGTCCTCAACAGAGCTGATACAGAGATGGATTGTTA 560
    |||||
Db 204 rleuglyserAlaTyHialAlaIargargaspMetleuaspAlaLeuGlyIleThrAl 224
    ::::|
QY 561 TGTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTTATCCCGAGTCTCATTTCT 620
    ::::|
Db 224 aleuLeuAanValSerSerAspCysProAen--HisPheGluGlyHisTyrgIntYrly 243
    ::::|
QY 621 GCCTGTCCTGTGAAATGACACGCTTTTGTGTAGAAATTTTGCCTGGTTGACAAATCAGT 680
    ::::|
Db 243 sCysIleProValGluaspAsnHisIysAlaAspIleSerSerTrpPheMetGluAlaI 263
    ::::|
QY 681 AGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTTCTTAGTCACTGTTAGCTG 740
    ::::|
Db 263 eglIntYrIleaspAlaValIysaspCysargGlyIargValIleuValHisCysGlnAlaG 283
    ::::|
QY 741 GATCTCCGCTCCGCCACCATCGCTATCGCTTACATCATGAAGAGATGAGCATGTCTTT 800
    ::::|
Db 283 yIleSerArgSerAlaThrIleCysleuAlaTyrlleuMetMetIysIysargValArgle 303
    ::::|
QY 801 AGATGAAGCTTACAGATTGTGTGAAGAAAAAGACCTACTATATCTCCAAACTTCAATT 860
    ::::|
Db 303 uGluGluAlaPheGluPheValIysGlnIargSerIleIleSerProAsnPheserPh 323
    ::::|
QY 861 TCTGGGCCAACTCTGACTATGAGAAAGAAATT 894
    ::::|
Db 323 emetGlyGlnleuGlnPheGlnSerGlnVal 334
    ::::|
```

Search completed: June 21, 2004, 12:43:28
Job time : 17.5553 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 12:17:10 ; Search time 5703.02 Seconds

(without alignments)
6885.611 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 906

Sequence: 1 atgcccacagatgatggtg.....agaagataaagaccagact 906

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ph:*
8: gb_ph:*
9: gb_ph:*
10: gb_ph:*
11: gb_ph:*
12: gb_ph:*
13: gb_ph:*
14: gb_ph:*
15: gb_ph:*
16: gb_ph:*
17: gb_ph:*
18: gb_ph:*
19: gb_ph:*
20: gb_ph:*
21: gb_ph:*
22: gb_ph:*
23: gb_ph:*
24: gb_ph:*
25: gb_ph:*
26: gb_ph:*
27: gb_ph:*
28: gb_ph:*
29: gb_ph:*
30: gb_ph:*
31: gb_ph:*
32: gb_ph:*
33: gb_ph:*
34: gb_ph:*
35: gb_ph:*
36: gb_ph:*
37: gb_ph:*
38: gb_ph:*
39: gb_ph:*
40: gb_ph:*
41: gb_ph:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	906	100.0	5450	6	AX482439
2	906	100.0	5450	6	AX482478
3	902.8	99.6	1998	6	AX260342
4	902.8	99.6	2071	6	AX921917
5	902.8	99.6	2118	6	AX099933
6	902.8	99.6	2732	6	AX180875
7	902.8	99.6	3059	6	AX278451
8	902.8	99.6	3284	6	BC042101
9	902.8	99.6	3496	6	AX441210
10	902.8	99.6	3521	6	AB052156
11	902.8	99.6	3544	6	AX260340
12	902.8	99.6	3566	6	AF506796
13	902.8	99.6	3766	6	AX374994
14	902.8	99.6	4790	6	BD171157
15	902.8	99.6	4790	6	BD183422
16	902.8	99.6	4790	6	AB051487
17	901.2	99.5	2102	6	AX713989
18	901.2	99.5	2102	9	AK055973
19	901.2	99.5	3104	6	AX405700
20	901.2	99.5	5111	6	AX482372
21	817.8	90.3	2200	6	AX921915
22	797.2	88.0	4874	10	AB052157
23	797.2	88.0	4943	10	BC059232
24	797.2	88.0	4975	10	BC057321
25	794	87.6	4992	10	AF345951
26	782.8	86.4	2756	6	AX482444
27	756.8	83.5	2807	9	BC031643
28	727.6	80.3	749	6	AX868469
29	727.6	80.3	749	6	BD148531
30	714.2	78.8	4026	10	AF345953
31	564.8	62.3	1935	9	AY038927
32	564.8	62.3	3332	6	AX441229
33	463.6	51.2	4837	10	AF345952
34	404.8	44.7	3231	5	BC063941
35	383.8	42.4	3661	10	AF345954
36	365.8	40.4	2377	6	AR255944
37	365.8	40.4	2377	9	HS027193
38	364.2	40.2	2264	9	BC045110
39	363.6	40.1	2453	6	A59888
40	363.6	40.1	2453	10	MM0771GN
41	363.6	40.1	2600	10	BC052705
42	351	38.7	2415	6	A59887
43	335.8	37.1	2476	6	AX552119
44	317	35.0	408	6	AX482446
45	311.6	34.4	183797	9	AC022400

ALIGNMENTS

RESULT 1
LOCUS AX482439 5450 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 108 from Patent WO02057460.
ACCESSION AX482439
VERSION AX482439.1 GI:22316984
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Todderud,C.G., Bol,D., Finger,J., Lee,J., Nelson,T., Schieven,G.,
Suchard,S., Banae,D., Bassolino,D., Feder,J., Krystek,S.,
Mcatee,P., Muttler,G., Siemers,N., Jackson,D.G. and Ramanathan,C.

TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 108 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)
FEATURES Location/Qualifiers
source 1..5450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538..2535
/note="unnamed protein product"

CDS

/codon_start=1
/protein_id="CAD44451.1"
/db_xref="GI:22316985"
/translation="MAHEMIGTOIVTERLVALLSESTKRVLLIDSRPREVYNTSHILE
AININCSKLMRRLOQDKVITELIIOHSAHKVDIDSCQVYVYDSSODVASISDC
FLVYLGLKLEKSPNSVHLIAGPAEPRCPGLCEKSTLVPTGISQCLPVANIGPT
RILPNLYGCGORDVANKELIQNGIGVYVNAATYCPKDPFIPESHFLRPVNDSECK
KRPITSPNPNPFIKAKASNGCVLHCLAGISATLALAYIKRMDSLDPAVFEK
TRLSPPCADATSEAGQRPVHPASVPSVPSLIDSPILVQALSGHLNADRLD
SNKTKRFSFDIKSVSASMAASLHGFSSSDALEYKPSITLDGNTKICQPSVPE
LSQETPETSDEKASIPKRLQTPARPSOSKRLHSVTSSTGASLSLPLHRSQS
VEDNYHTSFLFGLSTSQHLLTKSAGLGLKQMSDILAQTSIPSLTSMYFATESHP
YSASAIYGSASYSAYSCSLPTCGDQVYVRRROKPSDRADSRMSWHEESPKEQPK
RSCQMEFESIMSENRSREELGKVGSSPSGSMETIEVS"

ORIGIN

Query Match 100.0%; Score 906; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 5,4e-268;
Matches 906; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN
1 ATGGCCCATGAGATGATTGAACTCAATTGTACTGAGAGTTGGTGGCTCTGCTGAA 60
538 ATGGCCCATGAGATGATTGAACTCAATTGTACTGAGAGTTGGTGGCTCTGCTGAA 60
61 AGTGAACGAAAAAGTCTGCTAATTGATAGCCGGCCATTGTGGAATACATACATCC 120
538 AGTGAACGAAAAAGTCTGCTAATTGATAGCCGGCCATTGTGGAATACATACATCC 120
121 CACATTTGGAGCCATTATATCAACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 180
658 CACATTTGGAGCCATTATATCAACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 180
181 GACAAAGTGTAAATTACAGAGCTCATCCGACCATTCAGCGAAACATAGATTGATGAT 240
718 GACAAAGTGTAAATTACAGAGCTCATCCGACCATTCAGCGAAACATAGATTGATGAT 240
241 TGCAGTGAAGGTTGATTACATCAATCAAGCTCCAAAGATGTGGCTCTCTCTTCA 300
778 TGCAGTGAAGGTTGATTACATCAAGCTCCAAAGATGTGGCTCTCTCTCTTCA 300
301 GACTGTTTTCTACTGTACTTCTGGGTTAACTGAGAAAGAGCTTCAACTGTCTCACTG 360
838 GACTGTTTTCTACTGTACTTCTGGGTTAACTGAGAAAGAGCTTCAACTGTCTCACTG 360
361 CTTCGAGGAGGTTTGCAGATTTCTCTGTTGTTCCCTGGCTCTGTAAGAAATTC 420
898 CTTCGAGGAGGTTTGCAGATTTCTCTGTTGTTCCCTGGCTCTGTAAGAAATTC 420
421 ACTTAAGTCCCTACCTGATTTCTCAGCTTGGCTTACCTGTTGCCAACTTGGGCAAC 480
958 ACTTAAGTCCCTACCTGATTTCTCAGCTTGGCTTACCTGTTGCCAACTTGGGCAAC 480
481 CGAATTCCTCCCAATCTTTATCTTGGCTGCGACGAGATGTCCTCAACAGAGACTGATA 540
1018 CGAATTCCTCCCAATCTTTATCTTGGCTGCGACGAGATGTCCTCAACAGAGACTGATA 540
541 CAGAGAAATGGATGGTATGTTAAATGACAGTATACCTGCTCCAAAGCTGACTT 600
1078 CAGAGAAATGGATGGTATGTTAAATGACAGTATACCTGCTCCAAAGCTGACTT 600
601 ATCCCGAGTCTCATTTCTCGGTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 660

Db 1138 ATCCCGAGTCTCATTTCTCGGTGTGCTGTGAATGACCTTTGTGAGAAATTTTG 1197
Qy 661 CCGTGGTTGGAACAATCAGTATGATTGATGAAAGCAAAAGCCTCAATGATGTGTT 720
Db 1198 CCGTGGTTGGAACAATCAGTATGATTGATGAAAGCAAAAGCCTCAATGATGTGTT 720
Qy 721 CTAGTGCATGTTTACCTGGGATCTCCGCTCCGCAACATCGCTATCGCTACATCATG 780
Db 1258 CTAGTGCATGTTTACCTGGGATCTCCGCTCCGCAACATCGCTATCGCTACATCATG 780
Qy 781 AAGAGATGACATGCTTTGATGAGCTTACAGATTTGTGAAAGAAAAAGACCTACT 840
Db 1318 AAGAGATGACATGCTTTGATGAGCTTACAGATTTGTGAAAGAAAAAGACCTACT 840
Qy 841 ATATGTCGAACCTCAATTTTCTGGGCAACTCTGGAATGAGAAAGATTAAAGAC 900
Db 1378 ATATGTCGAACCTCAATTTTCTGGGCAACTCTGGAATGAGAAAGATTAAAGAC 900
Qy 901 CAGACT 906
Db 1438 CAGACT 1443

RESULT 2
AX482478
LOCUS AX482478 5450 bp DNA linear 16-AUG-2002
DEFINITION Sequence 147 from Patent WO02057460.
ACCESSION AX482478
VERSION AX482478.1 GI:22316999
KEYWORDS

SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Toddlerud, C.G., Bol, D., Finger, J., Lee, L., Nelson, T., Schieven, G.,
Mcatee, P., Mutter, G., Siemers, N., Jackson, D.G., and Ramanathan, C.

TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 147 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)

FEATURES

source 1..5450
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538..2535
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD44451.1"
/db_xref="GI:22317000"

CDS

/translation="MAHEMIGTOIVTERLVALLSESTKRVLLIDSRPREVYNTSHILE
AININCSKLMRRLOQDKVITELIIOHSAHKVDIDSCQVYVYDSSODVASISDC
FLVYLGLKLEKSPNSVHLIAGPAEPRCPGLCEKSTLVPTGISQCLPVANIGPT
RILPNLYGCGORDVANKELIQNGIGVYVNAATYCPKDPFIPESHFLRPVNDSECK
KRPITSPNPNPFIKAKASNGCVLHCLAGISATLALAYIKRMDSLDPAVFEK
TRLSPPCADATSEAGQRPVHPASVPSVPSLIDSPILVQALSGHLNADRLD
SNKTKRFSFDIKSVSASMAASLHGFSSSDALEYKPSITLDGNTKICQPSVPE
LSQETPETSDEKASIPKRLQTPARPSOSKRLHSVTSSTGASLSLPLHRSQS
VEDNYHTSFLFGLSTSQHLLTKSAGLGLKQMSDILAQTSIPSLTSMYFATESHP
YSASAIYGSASYSAYSCSLPTCGDQVYVRRROKPSDRADSRMSWHEESPKEQPK
RSCQMEFESIMSENRSREELGKVGSSPSGSMETIEVS"

ORIGIN

Query Match 100.0%; Score 906; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 5,4e-268;
Matches 906; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN
1 ATGGCCCATGAGATGATTGAACTCAATTGTACTGAGAGTTGGTGGCTCTGCTGAA 60

538 ATGGCCCATGAGATGATGGAAGTCAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGAA 597
QY 61 AGTGAACCGAAAAAGTCTGCTAATTTGATAGCCGGCATTGTGGAAATACATATCC 120
DB 598 AGTGAACCGAAAAAGTCTGCTAATTTGATAGCCGGCATTGTGGAAATACATATCC 657
QY 121 CACATTTGGGAACCATTAATATCACTGCTCAAGCTTATGAAAGGAAGTTGCAACG 180
DB 658 CACATTTGGGAACCATTAATATCACTGCTCAAGCTTATGAAAGGAAGTTGCAACG 717
QY 181 GACAAAGTTAATTAACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATGAT 240
DB 718 GACAAAGTTAATTAACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATGAT 777
QY 241 TGCAGTCAGAGGTTGATTTACGATCAAAAGTCCCAAGATGTTGCTCTCTCTTCA 300
DB 778 TGCAGTCAGAGGTTGATTTACGATCAAAAGTCCCAAGATGTTGCTCTCTCTTCA 837
QY 301 GACTGTTTCTCACTGATCTTGGGTTAACTGGAAGAGGTTCACTGTTCACTG 360
DB 838 GACTGTTTCTCACTGATCTTGGGTTAACTGGAAGAGGTTCACTGTTCACTG 897
QY 361 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTTCCCTGGCTCTGTGAAGAAATCC 420
DB 898 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTTCCCTGGCTCTGTGAAGAAATCC 957
QY 421 ACTTAGTCCCTAAGCTGATTTCTCAGCTTGTCTTACCTGTTCCAAATTTGGCCAAAC 480
DB 958 ACTTAGTCCCTAAGCTGATTTCTCAGCTTGTCTTACCTGTTCCAAATTTGGCCAAAC 1017
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCAAGCGAAGTCTCTCAAGAGGCTGATA 540
DB 1018 CGAATTTCTCCCAATCTTTATCTTGGCTGCAAGCGAAGTCTCTCAAGAGGCTGATA 1077
QY 541 CAGCAAAATGGAGTTGGTTATGTTAAATGACAGCTATACCTGCAAAAGCTGACTT 600
DB 1078 CAGCAAAATGGAGTTGGTTATGTTAAATGACAGCTATACCTGCAAAAGCTGACTT 1137
QY 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG 660
DB 1138 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG 1197
QY 661 CCGTGGTGGACAAATCAGTAAATTTTATGAAAGAAAGCAAGCTTGAATGATGTT 720
DB 1198 CCGTGGTGGACAAATCAGTAAATTTTATGAAAGAAAGCAAGCTTGAATGATGTT 1257
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATG 780
DB 1258 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATG 1317
QY 781 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAGAAAAGACCTTACT 840
DB 1318 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAGAAAAGACCTTACT 1377
QY 841 ATATCTCCCAATCTTCAATTTTCTGGGCAACTCTGGAATGAAAGAAAGATTTAAGAC 900
DB 1378 ATATCTCCCAATCTTCAATTTTCTGGGCAACTCTGGAATGAAAGAAAGATTTAAGAC 1437
QY 901 CAGACT 906
DB 1438 CAGACT 1443

RESULT 3
AX260342 1998 bp DNA linear PAT 26-OCT-2001
LOCUS Sequence 3 from Patent WO0173059.
ACCESSION AX260342
VERSION AX260342.1 GI:16509305
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS
TITLE
Meyers, R.A.
3862 and 2117: dual specificity phosphatase molecules and uses
therefor
JOURNAL
Patent: WO 0173059-A 3 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1.1998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.6%; Score 902.8; DB 6; Length 1998;
Best Local Similarity 99.8%; Pred. No. 4,6e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATGGAAGTCAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGAA 60
DB 1 ATGGCCCATGAGATGATGGAAGTCAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGAA 60
QY 61 AGTGAACCGAAAAAGTCTGCTAATTTGATAGCCGGCATTGTGGAAATACATATCC 120
DB 61 AGTGAACCGAAAAAGTCTGCTAATTTGATAGCCGGCATTGTGGAAATACATATCC 120
QY 121 CACATTTGGGAACCATTAATATCACTGCTCAAGCTTATGAAAGGAAGTTGCAACG 180
DB 121 CACATTTGGGAACCATTAATATCACTGCTCAAGCTTATGAAAGGAAGTTGCAACG 180
QY 181 GACAAAGTTAATTAACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATGAT 240
DB 181 GACAAAGTTAATTAACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATGAT 240
QY 241 TGCAGTCAGAGGTTGATTTACGATCAAAAGTCCCAAGATGTTGCTCTCTCTTCA 300
DB 241 TGCAGTCAGAGGTTGATTTACGATCAAAAGTCCCAAGATGTTGCTCTCTCTTCA 300
QY 301 GACTGTTTCTCACTGATCTTGGGTTAACTGGAAGAGGTTCACTGTTCACTG 360
DB 301 GACTGTTTCTCACTGATCTTGGGTTAACTGGAAGAGGTTCACTGTTCACTG 360
QY 361 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTTCCCTGGCTCTGTGAAGAAATCC 420
DB 361 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTTCCCTGGCTCTGTGAAGAAATCC 420
QY 421 ACTTAGTCCCTAAGCTGATTTCTCAGCTTGTCTTACCTGTTCCAAATTTGGCCAAAC 480
DB 421 ACTTAGTCCCTAAGCTGATTTCTCAGCTTGTCTTACCTGTTCCAAATTTGGCCAAAC 480
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCAAGCGAAGTCTCTCAAGAGGCTGATA 540
DB 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCAAGCGAAGTCTCTCAAGAGGCTGATA 540
QY 541 CAGCAAAATGGAGTTGGTTATGTTAAATGACAGCTTATACCTGCAAAAGCTGACTT 600
DB 541 CAGCAAAATGGAGTTGGTTATGTTAAATGACAGCTTATACCTGCAAAAGCTGACTT 600
QY 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG 660
DB 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG 660
QY 661 CCGTGGTGGACAAATCAGTAAATTTTATGAAAGCAAGCTTGAATGATGTT 720
DB 661 CCGTGGTGGACAAATCAGTAAATTTTATGAAAGCAAGCTTGAATGATGTT 720
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCCACCATCGCTACATCATG 780
DB 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCCACCATCGCTACATCATG 780
QY 781 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAGAAAAGACCTTACT 840
DB 781 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAGAAAAGACCTTACT 840

QY 841 ATATCTCCAACTTCAATTTCTGGCCCACTCTGAGCTATGAGAAGATTAGAAC 900
 Db 841 ATATCTCCAACTTCAATTTCTGGCCCACTCTGAGCTATGAGAAGATTAGAAC 900
 QY 901 CAGACT 906
 Db 901 CAGACT 906

RESULT 4
 AX921917 2071 bp DNA linear PAT 18-DEC-2003
 LOCUS Sequence 257 from Patent WO02068649.
 ACCESSION AX921917
 VERSION AX921917.1 GI:40215410
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 Patent: WO 02068649-A 257 06-SEP-2002;
 AUTHORS Curagen Corporation (US)
 JOURNAL Location/Qualifiers
 FEATURES
 source 1.
 1.2071
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 99.6%; Score 902.8; DB 6; Length 2071;
 Best Local Similarity 99.8%; Pred. No. 4.6e-267;
 Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGGCCCATGAGATGATGAACTCAATTGTTACTGAGGTGGCTGCTGGA 60
 Db 61 ATGGCCCATGAGATGATGAACTCAATTGTTACTGAGGTGGCTGCTGGA 60
 QY 61 AGTGAACGGAAGAAAGTCTGCTTAATGATGAGCCGCAATTTGGGATACATCATCC 120
 Db 121 AGTGAACGGAAGAAAGTCTGCTTAATGATGAGCCGCAATTTGGGATACATCATCC 120
 QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 180
 Db 121 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 180
 QY 181 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 240
 Db 181 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 240
 QY 241 TGCAGTCAGAAAGTGTAGTTACGATCAAGCTCCCAAGATGTGCTCTCTCTCA 300
 Db 241 TGCAGTCAGAAAGTGTAGTTACGATCAAGCTCCCAAGATGTGCTCTCTCTCA 300
 QY 301 GACTGTTTCTCAGTCTCTGCGGTAACTGGAAGAGCTTCAACTGTGTCACCTG 360
 Db 301 GACTGTTTCTCAGTCTCTGCGGTAACTGGAAGAGCTTCAACTGTGTCACCTG 360
 QY 361 GACTGTTTCTCAGTCTCTGCGGTAACTGGAAGAGCTTCAACTGTGTCACCTG 420
 Db 361 GACTGTTTCTCAGTCTCTGCGGTAACTGGAAGAGCTTCAACTGTGTCACCTG 420
 QY 421 ACTTAAGTCCCTACCTGATTTCTCGAGCTTCTGCTGTTCCCTGCGCTGGAAGAAATCC 480
 Db 421 ACTTAAGTCCCTACCTGATTTCTCGAGCTTCTGCTGTTCCCTGCGCTGGAAGAAATCC 480
 QY 481 ACTTAAGTCCCTACCTGATTTCTCGAGCTTCTGCTGTTCCCTGCGCTGGAAGAAATCC 540
 Db 481 ACTTAAGTCCCTACCTGATTTCTCGAGCTTCTGCTGTTCCCTGCGCTGGAAGAAATCC 540
 QY 541 CAGATTTTCCCAATTTTATCTGGCTGCGAGAGATGCTTCAACAAAGAGCTGATG 600
 Db 541 CAGATTTTCCCAATTTTATCTGGCTGCGAGAGATGCTTCAACAAAGAGCTGATG 600
 QY 600 CAGATTTTCCCAATTTTATCTGGCTGCGAGAGATGCTTCAACAAAGAGCTGATG 600
 Db 600 CAGATTTTCCCAATTTTATCTGGCTGCGAGAGATGCTTCAACAAAGAGCTGATG 600

Db 601 CAGCAGATGAGATGATGTTATGTTAAATGCAACATACCTGTCACAAACCTGACTTT 660
 QY 601 ATCCCGAGTCTCATTTCCGAGTGTGCTGTAATGACAGCTTTTGTGAAAAATTTTG 660
 Db 661 ATCCCGAGTCTCATTTCCGAGTGTGCTGTAATGACAGCTTTTGTGAAAAATTTTG 720
 QY 661 CCGTGTGTGAGCAAAATCAGTATGATTTCAATTGAGAAACAAAGCTTCAATGATGTT 720
 Db 721 CCGTGTGTGAGCAAAATCAGTATGATTTCAATTGAGAAACAAAGCTTCAATGATGTT 720
 QY 721 CTAATGCACTGTTTATGCTGAGTCTCCCGCTCCGCAACATGCTTATGATGATG 780
 Db 781 CTAATGCACTGTTTATGCTGAGTCTCCCGCTCCGCAACATGCTTATGATGATG 780
 QY 781 AAGAGATGAGCATGCTTTATGATGAAAGCTTACAGATTTGTGAAAAAGAAAGACTACT 840
 Db 841 AAGAGATGAGCATGCTTTATGATGAAAGCTTACAGATTTGTGAAAAAGAAAGACTACT 840
 QY 841 ATATCTCCAACTTCAATTTCTGGCCCACTCTGAGCTATGAGAAGATTAGAAC 900
 Db 901 ATATCTCCAACTTCAATTTCTGGCCCACTCTGAGCTATGAGAAGATTAGAAC 900
 QY 901 CAGACT 906
 Db 961 CAGACT 966

RESULT 5
 AX099933 2118 bp DNA linear PAT 02-APR-2001
 LOCUS Sequence 15 from Patent WO0120004.
 ACCESSION AX099933
 VERSION AX099933.1 GI:13538943
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 Yue, H., Tang, Y. T., Bandman, O., Hillman, D. L., Baughn, M. R.,
 Azimzal, Y., and Lu, D. A.
 TITLE Protein phosphatase and kinase proteins
 JOURNAL Patent: WO 0120004-A 15 22-MAR-2001;
 INCYTE Genomics, Inc. (US)
 FEATURES
 source 1.
 1.2118
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID NO: 1234795CB1"

ORIGIN

Query Match 99.6%; Score 902.8; DB 6; Length 2118;
 Best Local Similarity 99.8%; Pred. No. 4.6e-267;
 Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGGCCCATGAGATGATGAACTCAATTGTTACTGAGGTGGCTGCTGGA 60
 Db 25 ATGGCCCATGAGATGATGAACTCAATTGTTACTGAGGTGGCTGCTGGA 60
 QY 61 AGTGAACGGAAGAAAGTCTGCTTAATGATGAGCCGCAATTTGGGATACATCATCC 120
 Db 85 AGTGAACGGAAGAAAGTCTGCTTAATGATGAGCCGCAATTTGGGATACATCATCC 144
 QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 180
 Db 145 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 204
 QY 181 GACAAAGTGTATTAACAGAGCTTCCAGATTCAGCGAAGCAATAGGTTGACATTGAT 240
 Db 205 GACAAAGTGTATTAACAGAGCTTCCAGATTCAGCGAAGCAATAGGTTGACATTGAT 264

QY 241 TGCAGTCAGAGGTTGATTTAGATCAAGCTCCAGATGTTGCTCTCTCTCA 300
DB 265 TGCAGTCAGAGGTTGATTTAGATCAAGCTCCAGATGTTGCTCTCTCTCA 324
QY 301 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAGAGCTTCACTCTGACCTG 360
DB 325 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAGAGCTTCACTCTGACCTG 384
QY 361 CTTCAGAGTGGGTTTCTGAGTCTTCTGCTGTTTCTCTGCTCTCTGAGAGAAATTC 420
DB 385 CTTCAGAGTGGGTTTCTGAGTCTTCTGCTGTTTCTCTGCTCTCTGAGAGAAATTC 444
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCTTGTCTTACCTGTGCAATTTGGGCAACC 480
DB 445 ACTCTAGTCCCTACCTGATTTCTCAGCTTGTCTTACCTGTGCAATTTGGGCAACC 504
QY 481 CGAATTTCTCCCAATCTTTATCTTGTGCTGCGCAGAGATGCTCCAAAGAGCTGATA 540
DB 505 CGAATTTCTCCCAATCTTTATCTTGTGCTGCGCAGAGATGCTCCAAAGAGCTGATA 564
QY 541 CAGCAGATGGGATGGTTATGTGTAAATGCCAGCTTACCTGTCCAAAGCTGACTTT 600
DB 565 CAGCAGATGGGATGGTTATGTGTAAATGCCAGCTTACCTGTCCAAAGCTGACTTT 624
QY 601 ATCCCGAGTCTATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 625 ATCCCGAGTCTATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684
QY 661 CCGTGTGTCAGCAATCAGTATGATTTGATGAGAGCAAAAGCTTCAATGATGTTT 720
DB 685 CCGTGTGTCAGCAATCAGTATGATTTGATGAGAGCAAAAGCTTCAATGATGTTT 744
QY 721 CTAGTCAGCTGTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 745 CTAGTCAGCTGTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
QY 781 AAGAGATGAGCAATGCTTTATGATGAGCTTACGATTTGAGAGAGAGAGAGCTTACT 840
DB 805 AAGAGATGAGCAATGCTTTATGATGAGCTTACGATTTGAGAGAGAGAGAGCTTACT 864
QY 841 ATATCTCCAACTTCAATTTTCTGCGGCAACTCTCTGAGCTATGAGAGAGATTAAGAC 900
DB 865 ATATCTCCAACTTCAATTTTCTGCGGCAACTCTCTGAGCTATGAGAGAGATTAAGAC 924
QY 901 CAGACT 906
DB 925 CAGACT 930

RESULT 6
AX180875 2732 bp DNA linear PAT 06-AUG-2001
LOCUS
DEFINITION
ACCESSION
VERSION
AX180875.1 GI:15132703
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 Plozman, G.D., Martinez, R., Whyte, D., Manning, G., Sudarsanam, S.,
Hill, R.V. and Flanagan, P.
TITLE
JOURNAL
Mammalian protein phosphatases
Patent: WO 0146394-A 2 28-JUN-2001;
Sugen, Inc. (US)
FEATURES
source
1..2732
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.6%; Score 902.8; DB 6; Length 2732;
Best Local Similarity 99.6%; Pred. No. 4, 8e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
DB 538 ATGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 597
QY 61 AGTGAACGAGAAAGTGTCTGATTAATGATGATGATGATGATGATGATGATGATGAT 120
DB 598 AGTGAACGAGAAAGTGTCTGATTAATGATGATGATGATGATGATGATGATGATGAT 657
QY 121 CACATTTGGAGACCAATTAATCACTGCTCCAGCTTATGAGAGAGAGAGAGAGAGAG 180
DB 658 CACATTTGGAGACCAATTAATCACTGCTCCAGCTTATGAGAGAGAGAGAGAGAGAG 717
QY 181 GACAAAGTGTAAATTAAGAGCTATCAGAGCTATCAGAGCTATCAGAGCTATCAGAGCT 240
DB 718 GACAAAGTGTAAATTAAGAGCTATCAGAGCTATCAGAGCTATCAGAGCTATCAGAGCT 777
QY 241 TGCAGTCAGAGGTTGATTTAGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
DB 778 TGCAGTCAGAGGTTGATTTAGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 837
QY 301 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAGAGCTTCACTGTTCACTG 360
DB 838 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAGAGCTTCACTGTTCACTG 897
QY 361 CTTCAGAGTGGGTTTCTGAGTCTTCTGCTGTTTCTCTGCTGCTGCTGCTGCTGCTGCT 420
DB 898 CTTCAGAGTGGGTTTCTGAGTCTTCTGCTGTTTCTCTGCTGCTGCTGCTGCTGCTGCT 957
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCTTGTCTTACCTGTGCAATTTGGGCAACC 480
DB 958 ACTCTAGTCCCTACCTGATTTCTCAGCTTGTCTTACCTGTGCAATTTGGGCAACC 1017
QY 481 CGAATTTCTCCCAATCTTTATCTTGTGCTGCGCAGAGATGCTTCCAAAGAGCTGATA 540
DB 1018 CGAATTTCTCCCAATCTTTATCTTGTGCTGCGCAGAGATGCTTCCAAAGAGCTGATA 1077
QY 541 CAGCAGATGGGATGGTTATGTGTAAATGCGAGCTATACCTGTCCAAAGCTGACTTT 600
DB 1078 CAGCAGATGGGATGGTTATGTGTAAATGCGAGCTATACCTGTCCAAAGCTGACTTT 1137
QY 601 ATCCCGAGTCTATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 1138 ATCCCGAGTCTATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
QY 661 CCGTGTGTCAGCAATCAGTATGATTTCAATGAGAGCAAAAGCTTCAATGATGTTT 720
DB 1198 CCGTGTGTCAGCAATCAGTATGATTTCAATGAGAGCAAAAGCTTCAATGATGTTT 1257
QY 721 CTAGTCAGCTGTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 1258 CTAGTCAGCTGTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317
QY 781 AAGAGATGAGCAATGCTTTATGATGAGCTTACGATTTGAGAGAGAGAGAGAGCTTACT 840
DB 1318 AAGAGATGAGCAATGCTTTATGATGAGCTTACGATTTGAGAGAGAGAGAGAGCTTACT 1377
QY 841 ATATCTCCAACTTCAATTTTCTGCGGCAACTCTCTGAGCTATGAGAGAGATTAAGAC 900
DB 1378 ATATCTCCAACTTCAATTTTCTGCGGCAACTCTCTGAGCTATGAGAGAGATTAAGAC 1437
QY 901 CAGACT 906
DB 1438 CAGACT 1443

RESULT 7
AX278461 3059 bp DNA linear PAT 02-NOV-2001
LOCUS
DEFINITION
Sequence 1 from Patent WO0177340.

ACCESSION	AX278461
VERSION	AX278461.1
KEYWORDS	GI:16605915
SOURCE	
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	Dueker, K.
AUTHORS	Identification of a dual specificity phosphatase: dusp-10
TITLE	Patent: WO 01/7340-A 1 18-OCT-2001;
JOURNAL	MERCK PATENT GmbH (DE)
FEATURES	Location/Qualifiers
source	1..3059

CDS 127.2124

ORIGIN

Query Match	99.6%	Score 902.8;	DB 6;	Length 3059;
Best Local Similarity	99.8%	Pred No. 4.9e-267;		
Matches 904; Conservative	0.0%			

QY	1	ATGGCCCATGATGATTTGGAACTGAAATTGTTACTGAGAGGTTGGTCTCTGCTGGAA	60
Db	127	ATGGCCCATGATGATTTGGAACTGAAATTGTTACTGAGAGGTTGGTCTCTGCTGGAA	186
QY	61	AGTGAACGAAAAGGCTGCTAAATTGATATAGCCGCCATTGTGGAAATACATATATCC	120
Db	187	AGTGAACGAAAAGGCTGCTAAATTGATATAGTATGATACCGCCATTGTGGAAATACATATATCC	246
QY	121	CACATTTGAAAGCCATTATATATCAACTGCTTCAAGCTTATGAAAGCGAAGTTGGCAAG	180
Db	247	CACATTTGAAAGCCATTATATATCAACTGCTTCAAGCTTATGAAAGCGAAGTTGGCAAG	306
QY	181	GACAAAGTTTAATTACAGAGCTCATCAGAGTTACAGGAAACATAAGGTTGACATTGAT	240
Db	307	GACAAAGTTTAATTACAGAGCTCATCAGAGTTACAGGAAACATAAGGTTGACATTGAT	366
QY	241	TGCATCGAAGGTTGATGATTAAGATCAAAAGCTCCCAAGATTTGCTCTCTCTTCA	300
Db	367	TGCATCGAAGGTTGATGATTAAGATCAAAAGCTCCCAAGATTTGCTCTCTCTCTTCA	426
QY	301	GACGTTTTTCTCATCTGTACTTCTGGGTAACCTGAGAAAGCTTCAACTCTGTTCACTG	360
Db	427	GACGTTTTTCTCATCTGTACTTCTGGGTAACCTGAGAAAGCTTCAACTCTGTTCACTG	486
QY	361	CTTGACAGTGGGTTTGCTGAGTTCTCTCGTTGTTTCTCGGCTCTGTGTGAAGAAATCC	420
Db	487	CTTGACAGTGGGTTTGCTGAGTTCTCTCGTTGTTTCTCGGCTCTGTGTGAAGAAATCC	546
QY	421	ACTCTAGTCCCTACCTGCAATTTCTACGCTTCTTACCTGTGCGCAACATTTGGGCGCAAC	480
Db	547	ACTCTAGTCCCTACCTGCAATTTCTAGCCTTCTTACCTGTGCGCAACATTTGGGCGCAAC	606

QY	481	CGAATTCCTCCCAATCTTTATCTTGCGCTGCAGCGAGATGTCTCTCAACAAGAGCTGATA	54
Db	607	CGAATTCCTCCCAATCTTTATCTTGCGCTGCAGCGAGATGTCTCTCAACAAGAGCTGATG	66
QY	541	CAGCAGAAATGGAGATGGTTATGCTTTAAATGCCAGCTATACCTGTCCAAAGCCTACCTT	600
Db	667	CAGCAGAAATGGAGATGGTTATGCTTTAAATGCCAGCAGATACTGTCTCCAAAGCCTACCTT	722
QY	601	ATCCCCGAGCTCAATTCCTGCGCTGCGCTGTGAATGACAGCTTTTGTGAGAAATTTTG	660
Db	727	ATCCCCGAGCTCAATTCCTGCGCTGCGCTGTGAATGACAGCTTTTGTGAGAAATTTTG	786
QY	661	CCGTGGTGGACAATACATAGATTTCATTTGAGAAAGCAAAAGCCTCCATGATGTGTT	720
Db	787	CCGTGGTGGACAATACATAGATTTCATTTGAGAAAGCAAAAGCCTCCATGATGTGTT	846
QY	721	CTAGTGCACGTTTAGCTGGGATCTCCCGCTCCGCCACATCGCTATGCGCTTACATCATG	780
Db	847	CTAGTGCACGTTTAGCTGGGATCTCCCGCTCCGCCACATCGCTATGCGCTTACATCATG	906
QY	781	AAGGAGATGACATGCTTTAGATGATGAAGCTTACAGATTGTGGAAGAAAAAGAACTTACT	840
Db	907	AAGGAGATGACATGCTTTAGATGATGAAGCTTACAGATTGTGGAAGAAAAAGAACTTACT	966
QY	841	ATAATCTCCAAACTTCATTTCTGGGCCCACTCTGSACTATGAGAAAGAAAGATTAAAGAC	900
Db	967	ATAATCTCCAAACTTCATTTCTGGGCCCACTCTGSACTATGAGAAAGAAAGATTAAAGAC	1026
QY	901	CAGACT 906	
Db	1027	CAGACT 1032	

RESULT #	LOCUS	DEFINITION	ACCESSION	VERSION	SOURCE	ORGANISM	REFERENCE	AUTHORS
BC042101	3284 bp	mRNA	linear	PRI 07-OCT-2003				
BC042101		Homo sapiens dual specificity phosphatase 16, mRNA (cDNA clone MGC:50665 IMAGE:440039), complete cds.						
BC042101								
BC042101.1	GI:27469788	MGC.						
		Homo sapiens (human)						
		Homo sapiens						
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
		Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Hopkins R.F., Zeeberg B., Bluet K.H., Schaefer C.F., Bat N.K., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Scheetz T.E., Brownstein M.B., Donaldson M.F., Casavant T.L., Carninci P., Brannstrom R.D., Mullany S.J., Bosak S.A., Loquellano N.A., Peters G.U., McKernan K.J., Malek U.A., Gunaratne P.H., Richards S., Villalón D.C., Hale S., Garcia A.M., Gay L.J., Huliyil S.W., Worley K.C., Hale S., Sodergren E.U., Lu X., Gibbs R.A., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Scherch A., Schein J.E., Jones S.J. and Wertz M.A.						
		Generation and initial analysis of more than 15,000 human and mouse cDNA sequences						
		Proc. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)						
		12477932						
		2 (bases 1 to 3284)						
		Strausberg, R.						
		Direct Submission						
		Submitted (23-DEC-2002) National Center for Biotechnology Information						

REMARK
COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) md@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 88 Row: a Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: GenomScan gene prediction.

FEATURES
source

Location/Qualifiers
1. .3284
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:50665 IMAGE:4400399"
/issue_type="Duodenum, adenocarcinoma"
/clone_lib="NIH MGC_88"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .3284
/gene="DUSP16"
/note="Synonyms: KIAA1700, MKP-7, MKP7"
/db_xref="LocusID:80824"
/db_xref="MIM:607175"

CDS

332..2320
/codon_start=1
/product="DUSP16 protein"
/protein_id="AAH42101.1"
/db_xref="GI:27469789"
/db_xref="LocusID:80824"
/translation="MAHBMICQIVTERLVALLSEKTVLLIDSPPEVYNTSHLE
AININSGKIMKRRILQODKVLITBLIHSKHKVIDCSQVYVYDSSQDVASLSDDC
PLTVLGLKESFNVSVHLAAGFAEFSRCPGLCEKSKTLVPTCLISQPLVANIIGPT
RLIPNLVLCORVNLKELMOQWIGVYVNASNTCPDPFIPESHRLRVVDSFCEK
ILPWLDSVDFIEKAKASNGCVLHCLAGISRSATIAIAYIMKRMMSLDEAVRFXE
KPTLIPNPNFPLGOLLIDYEKIKNOTGASGPKSLHLHEKNEPVPVAVSREGQKSE
TPLSPCADSATSEAGQRPVHPASVPSVQPSLESPLYOALSGHLASDLRLEBKN
LKRSFSLDIKSVISASMAASHLGFSSSDALRYKPTLLDSTNLCPFSPOELSE
QPTETSPDKREASIPKILQYARPSDSQSRHLHVSRTSSGTRSLSLPLHRSQVED
NYHTSFLPGLSTSQOHLTKYSAGLGLKMSDILAPQTSPLTSWYFATESHFFVSA
SAIYGSASYSAYSCSLPTCGDOVYSVRRORPDRADRSRSHBESPFEEKFKRS
COMEPGESIMSENRSEELGKVGSSSPFGSMELIEVS"
misc_feature
380..730
/note="RHOD; Region: Rhodanese Homology Domain"
/db_xref="CDD:smart00450"
803..1222
/note="DSPC; Region: Dual specificity phosphatase,
catalytic domain. Ser/Thr and Tyr protein phosphatases.
The enzyme's tertiary fold is highly similar to that of
tyrosine-specific phosphatases, except for a 'recognition'
region"
/db_xref="CDD:pfam00782"

ORIGIN

Query Match 99.6%; Score 902.8; DB 9; Length 3284;
Best Local Similarity 99.8%; Pred. No. 4.9e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 ATGGCCCATGAGATGATGGAAGTCAATTTGTTACTGAGAGTTGGTGGCTGTCTGAGAA 60
|||||

Db 332 ATGGCCCATGAGATGATGGAAGTCAATTTGTTACTGAGAGTTGGTGGCTGTCTGAGAA 391
QY 61 AGTGAACCGGAAAAAGTGTCTGCTAATTGATAGCGGGCCATTGTGGAATTCATATCC 120
Db 392 AGTGAACCGGAAAAAGTGTCTGCTAATTGATAGCGGGCCATTGTGGAATTCATATCC 451
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCCAAGCTTATGAGAGGAAGTTGCAACAG 180
Db 452 CACATTTTGAAGCCATTATATCACTGCTCCCAAGCTTATGAGAGGAAGTTGCAACAG 511
QY 181 GACAAAGTTTATATCAAGAGCTCATCGACATTCAGGAAACATAGTTGACATTGAT 240
Db 512 GACAAAGTTTATATCAAGAGCTCATCGACATTCAGGAAACATAGTTGACATTGAT 571
QY 241 TGCAGTCAGAGTTTATATCAAGTCCCAAGATTTGCTCTCTCTCA 300
Db 572 TGCAGTCAGAGTTTATATCAAGTCCCAAGATTTGCTCTCTCTCA 631
QY 301 GACTGTTTCTCAGCTACTCTCGGGTAAACTGAGAGAGCTTCAACTCTGTCACCTG 360
Db 632 GACTGTTTCTCAGCTACTCTCGGGTAAACTGAGAGAGCTTCAACTCTGTCACCTG 691
QY 361 CTTCGAGTGGGTTTGTCTGAGTCTCTGTTTCTGCTGAGCTCTGAGAGAAATCC 420
Db 692 CTTCGAGTGGGTTTGTCTGAGTCTCTGTTTCTGCTGAGCTCTGAGAGAAATCC 751
QY 421 ACTTATGTTCCCTACCTGCAATTTCTCAGCTTGTCTTACCTGTTGCAACATTTGGCCAA 480
Db 752 ACTTATGTTCCCTACCTGCAATTTCTCAGCTTGTCTTACCTGTTGCAACATTTGGCCAA 811
QY 481 CGAATCTTCCCAATCTTATCTGAGCTGCAAGAGAGATGCTCTCAACAGAGAGTGATA 540
Db 812 CGAATCTTCCCAATCTTATCTGAGCTGCAAGAGAGATGCTCTCAACAGAGAGTGATA 871
QY 541 CAGCAGATGGGATGGTATGTTAAATGCGAGCTATACCTGTCMAAGCTGACCTT 600
Db 872 CAGCAGATGGGATGGTATGTTAAATGCGAGCTATACCTGTCMAAGCTGACCTT 931
QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAAATGACAGCTTTGTGAGAAATTTTG 660
Db 932 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAAATGACAGCTTTGTGAGAAATTTTG 991
QY 661 CCGTGTGTGACAAATCAGTAGATTTTATTGAGAAAGCAAAAGCTTCAATGATGTGTT 720
Db 992 CCGTGTGTGACAAATCAGTAGATTTTATTGAGAAAGCAAAAGCTTCAATGATGTGTT 1051
QY 721 CTAGTCACTGTTAGCTGGATCTCCCGCTCGCCACCATGCTATCGCTACATCATG 780
Db 1052 CTAGTCACTGTTAGCTGGATCTCCCGCTCGCCACCATGCTATCGCTACATCATG 1111
QY 781 AAGAGATGACATGCTTTTATGATGAGCTTACAGATTTGTGAAAGAAAAAGACTACT 840
Db 1112 AAGAGATGACATGCTTTTATGATGAGCTTACAGATTTGTGAAAGAAAAAGACTACT 1171
QY 841 ATATCTCCAACTCTCAATTTTCTGGGCAACTCTGAGCATGAGAGAAAGATTAGAAG 900
Db 1172 ATATCTCCAACTCTCAATTTTCTGGGCAACTCTGAGCATGAGAGAAAGATTAGAAG 1231
QY 901 CAGACT 906
Db 1232 CAGACT 1237

RESULT 9
AX441210 3496 bp DNA linear PAT 28-JUN-2002
LOCUS AX441210
DEFINITION Sequence 1 from Patent WO0226997.
ACCESSION AX441210
VERSION AX441210.1 GI:21665766
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS
1 LucHE, R.M. and Wei, B.
TITLE
Dep-16 dual-specificity phosphatase
JOURNAL
Patient: WO 0226997-A 1 04-APR-2002;
Cepivir, Inc. (US)

FEATURES
Source
1. 3496
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity 99.6%; Score 902.8; DB 6; Length 3496;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 1 ATGCCCATGAGATGATTTGGACCAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA 60
Db 562 ATGCCCATGAGATGATTTGGACCAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA 621
Oy 61 AGTGAACGGAAGAAAGTCTGCTAATTTGATAGCCGCAATTTGTTGGAATTAATATATCC 120
Db 622 AGTGAACGGAAGAAAGTCTGCTAATTTGATAGCCGCAATTTGTTGGAATTAATATATCC 120
Oy 121 CAATTTTGAAGCCATTAATATATCAAGTCTGCAAGCTTATGAAAGCAAGTTGCAACAG 180
Db 682 CAATTTTGAAGCCATTAATATATCAAGTCTGCAAGCTTATGAAAGCAAGTTGCAACAG 180
Oy 181 GACAAAGTGTAAATACAGAGCTATCCAGCAATTCAGCAAGCAATTAAGTTGATATATAT 240
Db 742 GACAAAGTGTAAATACAGAGCTATCCAGCAATTCAGCAAGCAATTAAGTTGATATATAT 240
Oy 241 TGCAGTCAAGAGTTGATTTAGATCAAGCAAGTCCCAAGATTTGCTCTCTCTTCA 300
Db 802 TGCAGTCAAGAGTTGATTTAGATCAAGCAAGTCCCAAGATTTGCTCTCTCTTCA 300
Oy 301 GACTGTTTCTCACTGATCTTCTGGGTAACCTGAGAAAGCTTCAACTGTTGACCTG 360
Db 862 GACTGTTTCTCACTGATCTTCTGGGTAACCTGAGAAAGCTTCAACTGTTGACCTG 360
Oy 361 CTTCAGAGTGGTTTGGTGAATTTCTCTGTTGTTTCCCGGCTCTGTTGAAAGAAATCC 420
Db 922 CTTCAGAGTGGTTTGGTGAATTTCTCTGTTGTTTCCCGGCTCTGTTGAAAGAAATCC 420
Oy 421 ACTCTAGTCCCTACCTGATCTTCTGAGCTTCTGCTTACCTGTTGCCAATTTGGCCAAC 480
Db 982 ACTCTAGTCCCTACCTGATCTTCTGAGCTTCTGCTTACCTGTTGCCAATTTGGCCAAC 480
Oy 481 CGAATTTCTCCCAATTTTATCTTGGCTGCGCAGAGATGTCCTCAAGAGAGCTGATG 1041
Db 1042 CGAATTTCTCCCAATTTTATCTTGGCTGCGCAGAGATGTCCTCAAGAGAGCTGATG 1041
Oy 541 CAGCAAGATGGAGTTGTTATGTTAAATGCCGCTATACCTGTTCCAAAGCTGACTTT 600
Db 1102 CAGCAAGATGGAGTTGTTATGTTAAATGCCGCTATACCTGTTCCAAAGCTGACTTT 600
Oy 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGATGACGCTTTTGTGAGAAATTTTG 660
Db 1162 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGATGACGCTTTTGTGAGAAATTTTG 660
Oy 661 CCGTGTGTGAAGCAATGATGATTTGATTTGAAAGCAAGAAAGCTTCCAAATGATGTT 720
Db 1222 CCGTGTGTGAAGCAATGATGATTTGATTTGAAAGCAAGAAAGCTTCCAAATGATGTT 720
Oy 721 CTATGTCATGTTTATGTTGAGATCTCCGCTCTCCGCAACCATGCTATAGCTATATATG 780
Db 1282 CTATGTCATGTTTATGTTGAGATCTCCGCTCTCCGCAACCATGCTATAGCTATATATG 780
Oy 781 AAGAGATGAGCATGCTTTAGATGAGGTTACATTTGAAAGAAAGAAAGCACTACT 840
Db 1342 AAGAGATGAGCATGCTTTAGATGAGGTTACATTTGAAAGAAAGAAAGCACTACT 1401

```

Oy 841 ATATCTCAAACTGCAATTTTCTGGGCCCACTCTGAGTATGAGAGAAAGATTAGAAC 900
Db 1402 ATATCTCAAACTGCAATTTTCTGGGCCCACTCTGAGTATGAGAGAAAGATTAGAAC 1461
Oy 901 CAGACT 906
Db 1462 CAGACT 1467

RESULT 10

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 3521
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/cell_line="Jurkat"
1. 3521
/gene="MKP-7"
564. -2561
/gene="MKP-7"
/codon_start=1
/product="MAPK phosphatase-7"
/protein_id="BAB40814.1"
/db_xref="GI:13548677"
/translation="MAHEMTGTOIVTERLVALLSETEKYLIDSRPEVYNTSHILE
AATINCSKLMKRRLOODKVLITTELIGSAHAKVDIOSKVVVYDSSQDVASLSDC
FLTVLLGLKESFNSVHLIAGPFRSFRCPGLCEBKSTLVPTCLISQPLCPVANTGPT
ILPMLVYGCORVDLANKELMOQNGTGVNANSTCKPDPFIESHFLRVVDSRCEK
KRPTISPRNFTLGLDYERKIKNOTGASGPKSKLILHEKPNPVPAVSGGCKSE
TSLPSPCADSTISAAAGQRPVHPASVPSVQPSLIDSPVQLSLHLSADLIED
SNKRSFSLDIDISVSYSASMAASLHGFSSSDALEYKRPSTTLGNTKLCQFSPVDE
LSDPTPTSPDKERASIPKKLQTPARPSDOSRSLSVYRSSGTARSLPLRSGS
YSASAIYGSFLFGSLTSCQHLTSGAGLCTGMSDILAFQSTPSTLTSWYFATESSHF
VSAAIYGSFSAVSAYSCSLPTPCDOVYSVRKRPDRADSRSRWSHESPFKQPK
RFSQCMFEGSINSEMRSEBELGKTVSOSFSFSGSEILIEVS"

ORIGIN

Query Match

Best Local Similarity 99.6%; Score 902.8; DB 9; Length 3521;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 1 ATGCCCATGAGATGATTTGGACCAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA 60
Db 564 ATGCCCATGAGATGATTTGGACCAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA 623
Oy 61 AGTGAACGGAAGAAAGTCTGCTAATTTGATAGCCGCAATTTGTTGGAATTAATATATCC 120

```

Db 624 AGTGAAGGAAAGGCTGCTATTTGATGAGCGGCACTTTGTGGAATACATACATCC 663
Qy 121 CACATTTTGGAGGCATTAAATATCACTGCTCCAGCTTATGAAAGGAGTTGCAACAG 180
Db 684 CACATTTTGGAGGCATTAAATATCACTGCTCCAGCTTATGAAAGGAGTTGCAACAG 743
Qy 181 GACAAAGTGTAAATTAAGAGCTCATCAGCATTCAGGAAACATAAGTTGACATTGAT 240
Db 744 GACAAAGTGTAAATTAAGAGCTCATCAGCATTCAGGAAACATAAGTTGACATTGAT 803
Qy 241 TGGAGTCAGAGGTTGATGTTTACGATCAAGCTCCAGAGATGTTGCTCTCTCTTCA 300
Db 804 TGGAGTCAGAGGTTGATGTTTACGATCAAGCTCCAGAGATGTTGCTCTCTCTTCA 863
Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAAACTGGAGAGGCTTCAACTCTGCTCCTG 360
Db 864 GACTGTTTCTCACTGATCTTCTGGGTAAACTGGAGAGGCTTCAACTCTGCTCCTG 923
Qy 361 CTGGAGGTGGTTTGTGAGTCTCTGCTGTTTCCCTGCTGCTGTAAGGAAATCC 420
Db 924 CTGGAGGTGGTTTGTGAGTCTCTGCTGTTTCCCTGCTGCTGTAAGGAAATCC 983
Qy 421 ACTCTAGTCCCTTACCTGCACTTCTGAGCTTGTCACTGCTGCAACATTTGGGCCAAC 480
Db 984 ACTCTAGTCCCTTACCTGCACTTCTGAGCTTGTCTGCTGCTGCAACATTTGGGCCAAC 1043
Qy 481 CGAATTTCTCCCAATCTTATCTTGTGCTGCGCAGGAGATGCTCTCAAGAGAGCTGATA 540
Db 1044 CGAATTTCTCCCAATCTTATCTTGTGCTGCGCAGGAGATGCTCTCAAGAGAGCTGATA 1103
Qy 541 CAGCAGAGTGGAGTGTGTTATGTTAAATGCGAGCATATCCTGTCGAAGCTGACTTT 600
Db 1104 CAGCAGAGTGGAGTGTGTTATGTTAAATGCGAGCATATCCTGTCGAAGCTGACTTT 1163
Qy 601 ATCCCGAGTCTCATTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 1164 ATCCCGAGTCTCATTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1223
Qy 661 CCGTGTGGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 1224 CCGTGTGGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1283
Qy 721 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 1284 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1343
Qy 781 AAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 1344 AAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1403
Qy 841 ATATCTCAAACTCTCAATTTTCTGGGCAACTCTGCTGATGATGATGATGATGATGAT 900
Db 1404 ATATCTCAAACTCTCAATTTTCTGGGCAACTCTGCTGATGATGATGATGATGATGAT 1463
Qy 901 CAGACT 906
Db 1464 CAGACT 1469

RESULT 11
LOCUS AX260340 3544 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 1 from Patent WO0173059.
ACCESSION AX260340
VERSION AX260340.1 GI:16509303
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Meyers, R.A.

TITLE 38692 and 21117: dual specificity phosphatase molecules and uses
JOURNAL Patent: WO 0173059-A 1 04-OCT-2001.
Millemmium Pharmaceuticals, Inc. (US)
FEATURES
source location/Qualifiers
1..3544
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
589..2586
/note="unassigned protein product"
/protein_id="CAD10218.1"
/cdon_start=1
/db_xref="GI:16509304"
/db_xref="RefSeq:CAD10218.1"
/translation="MAHEMIGTIVTERVALLESTETKYLIDSPPEVYNTSHLE
AININCSKLMKRLQODKYLITELIGHAKHVDIDCSQKVVYDSDSDVALSDDC
FLTVLLGKLEKSFNSVHLAAGFAEBSRPGICEBSKTLVPCISQPCPLVANI
RILPMLYLCORDLNKEIMQONGIGYVLANASNTCPKPPFIPESHFLRVNDSPCEK
ILPWLKSDVDFIEKASNSCVLHCLAGISRSATIAIYIMRMMSIDEAYRKYK
KPTTSPNPNFLQQLIDYRKIKNOGASGRPSKTLHLKRENEVPVAVSEGCQSE
TPLEPCADUSATSEAGORPVHPASVPSVQPSLEDSPLVQALSGHLNLSADRLD
SNLKRKPSLIDIKSVYSASMAASLHFPSSBEDALRYKSTLIDGTNLKOPVQOE
LEQUTETSPPDEKASI PKLQTPARSDOSKRLHSVRTSSGTAQRLSLPLHRS
VEDNHTSFLPLGISTSOHLTKSAGLGMHSDILAPOTSTPLTSWYFATSSH
VSASATYGSASVAYSASOLPFCGPOVYSVRRROKDRADRSRMBHS PREKQFK
RNSCMFESBESIMSENRSRELEKVSQSSFSQSMETIEVS"

ORIGIN
Query Match 99.6%; Score 902.8; DB 6; Length 3544;
Best Local Similarity 99.8%; Pred. No. 5e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGCCCCATGAGATGATGAGTGAATCTCAATTTGTTACTGAGAGTGTGCTGCTGGA 60
Db 589 ATGCCCCATGAGATGATGAGTGAATCTCAATTTGTTACTGAGAGTGTGCTGCTGGA 648
Qy 61 AGTGAACGGAAGAAAGTGTGCTGCTAATGATGAGCGGCGCAATTTGTGAATCAATACATCC 120
Db 649 AGTGAACGGAAGAAAGTGTGCTGCTAATGATGAGCGGCGCAATTTGTGAATCAATACATCC 708
Qy 121 CACATTTTGGAGGCATTAAATATCACTGCTCCAGCTTATGAAAGGAGTTGCAACAG 180
Db 709 CACATTTTGGAGGCATTAAATATCACTGCTCCAGCTTATGAAAGGAGTTGCAACAG 768
Qy 181 GACAAAGTGTAAATTAAGAGCTCATCAGCATTCAGGAAACATAAGTTGACATTGAT 240
Db 769 GACAAAGTGTAAATTAAGAGCTCATCAGCATTCAGGAAACATAAGTTGACATTGAT 828
Qy 241 TGGAGTCAGAGGTTGATGTTTACGATCAAGCTCCAGAGATGTTGCTCTCTTCA 300
Db 829 TGGAGTCAGAGGTTGATGTTTACGATCAAGCTCCAGAGATGTTGCTCTCTTCA 888
Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAAACTGGAGAGGCTTCAACTCTGCTCCTG 360
Db 889 GACTGTTTCTCACTGATCTTCTGGGTAAACTGGAGAGGCTTCAACTCTGCTCCTG 948
Qy 361 CTGGAGGTGGTTTGTGAGTCTCTGCTGTTTCCCTGCTGCTGTAAGGAAATCC 420
Db 949 CTGGAGGTGGTTTGTGAGTCTCTGCTGTTTCCCTGCTGCTGTAAGGAAATCC 1008
Qy 421 ACTCTAGTCCCTTACCTGCACTTCTGAGCTTGTCACTGCTGCTGCTGCTGCTGCTGCTG 480
Db 1009 ACTCTAGTCCCTTACCTGCACTTCTGAGCTTGTCACTGCTGCTGCTGCTGCTGCTGCTG 1068
Qy 481 CGAATTTCTCCCAATCTTATCTTGTGCTGCGCAGGAGATGCTCTCAAGAGAGCTGATA 540
Db 1069 CGAATTTCTCCCAATCTTATCTTGTGCTGCGCAGGAGATGCTCTCAAGAGAGCTGATA 1128
Qy 541 CAGCAGAGTGGAGTGTGTTATGTTAAATGCGAGCTAATCTGCTGCAAGCTGACTTT 600
Db 1129 CAGCAGAGTGGAGTGTGTTATGTTAAATGCGAGCTAATCTGCTGCAAGCTGACTTT 1188

QY 601 ATCCCGAGCTTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db 1189 ATCCCGAGCTTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1248
QY 661 CCGTGGTTGACAAATCAGTAGATTTTCATTTGAGAAAGCAAAAGCTTCAATGATGTGT 720
Db 1249 CCGTGGTTGACAAATCAGTAGATTTTCATTTGAGAAAGCAAAAGCTTCAATGATGTGT 1308
QY 721 CTAGTGCACTGTTTGTGCGGATCTCCCGCTCCGCAACATGCTATGCTTACATCATG 780
Db 1309 CTAGTGCACTGTTTGTGCGGATCTCCCGCTCCGCAACATGCTATGCTTACATCATG 1368
QY 781 AAGAGATGACATGTCTTTAGATGAGAGCTTACAGATTTGTGAAAAGAAAAGACTTACT 840
Db 1369 AAGAGATGACATGTCTTTAGATGAGAGCTTACAGATTTGTGAAAAGAAAAGACTTACT 1428
QY 841 ATATCTCCAAATCTCAATTTTCTGGGCGCACTCTGACTATGAGAAAGATTAAGAC 900
Db 1429 ATATCTCCAAATCTCAATTTTCTGGGCGCACTCTGACTATGAGAAAGATTAAGAC 1488
QY 901 CAGACT 906
Db 1489 CAGACT 1494

RESULT 12
AF506796 3566 bp mRNA linear PRI 30-OCT-2003
LOCUS Homo sapiens dual specificity phosphatase 16 (DUSP16) mRNA,
AF506796 complete cds.
AF506796.1 GI:25573087

DEFINITION Homo sapiens
ACCESSION AF506796
VERSION AF506796.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Homo sapiens (human)
Homo sapiens (human)
MAPK phosphatase DUSP16/MKP-7, a candidate tumor suppressor for
Oncogene 22 (49), 7728-7736 (2003)
2 (bases 1 to 3566)
Hocmaert, I., Marynen, P. and Baens, M.
Submitted (26-APR-2002) Department for Human Genetics-Flanders
Interuniversity Institute for Biotechnology (VIB), Katholieke
Universiteit Leuven, Herestraat 49, Leuven B-3000, Belgium
Location/Qualifiers
1..3566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13-p12"
/feature_type="bone marrow"
1..3566
/gene="DUSP16"
/note="synonym: MKP7"
633..2630
/gene="DUSP16"
/EC_number="3.1.3.48"
/EC_number="3.1.3.16"
/codon_start=1
/product="dual specificity phosphatase 16"
/protein_id="AA075120.1"
/translation="MAHEMIGQIVTERIVALLSESTKYLIDSRPFVEYNTHILE
AIVNLSKLMKRLQDQVLTLEILIOHSAKHVVDIDQKVVVDOSODVASISDC
RIPLNLYGQORDVINKETMOONGIYVLANSNTCPKPDLPESHFRLVFNVSFCFK
ILPMLKSVDFIEKAKSNGCVLVHCLAGISRSATIAIAYIMKMDMSLDEAIFYKE

FEATURES
SOURCE

gene

CDS

ORIGIN
KRPISPNPFIQGLDYEKIKNOTGAGPKSLKILHLEKNEPVPVAVSBOGQSE
TPELSPCADSATSEAGQRPVHPASVSPVOPSLLEBPIVQALSGHLSDRLSD
SNLKSFEPLDIKSVSASMAALHGSSEEDLVEYKRSITLDTGNKLCQPSPOE
LSQUPERSPEKEASIPKILQTPRPSDOSKRLHSVRSSTAPRSLSPHRS
VDNHTSFLFGLSTQOHLTKSAGLKKMHDILAPQTSITSLSSVYFATESHF
YASAIYGSASVAYSQSLPTCGDQVYVRPRQRKPSDADRSRMBEESPFEQFK
RSCMEGESIMSERNREBELKVSGSSFSMSIIEVS"

Query Match 99.6%; Score 902.8; DB 9; Length 3566;
Best Local Similarity 99.8%; Pred. No. 5e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATGATGAACTCAATTTGTACTGAGAGGTGGTCTTCTGTGAA 60
Db 633 ATGGCCCATGATGATGATGAACTCAATTTGTACTGAGAGGTGGTCTTCTGTGAA 692
QY 61 AGTGAACGGAAAAAGCTGCTTAATTGATAGCCGGCAATTTGTGAATACATCATCC 120
Db 693 AGTGAACGGAAAAAGCTGCTTAATTGATAGCCGGCAATTTGTGAATACATCATCC 752
QY 121 CACATTTGAGACCATTAATATCACTGCTCAAGCTTATGAGCCGAGGTTGCAACG 180
Db 753 CACATTTGAGACCATTAATATCACTGCTCAAGCTTATGAGCCGAGGTTGCAACG 812
QY 181 GACAAAGTTAATTAACAGAGCTCATCCAGATTCAGGAAACATTAAGTTGACATGAT 240
Db 813 GACAAAGTTAATTAACAGAGCTCATCCAGATTCAGGAAACATTAAGTTGACATGAT 872
QY 241 TGCAGTCAGAAAGTTGATGATTAACATCAAGTCCCAAGATGTTGCTCTCTCTTCA 300
Db 873 TGCAGTCAGAAAGTTGATGATTAACATCAAGTCCCAAGATGTTGCTCTCTCTTCA 392
QY 301 GACTGTTTTCACACTGATCTCTGAGTAACTGAGAAAGCTTCAACTCTGTCACTG 360
Db 933 GACTGTTTTCACACTGATCTCTGAGTAACTGAGAAAGCTTCAACTCTGTCACTG 992
QY 361 CTGAGAGTGGGTGTCGAGTCTCTGTTGTTCCCTGAGCTCTGTGAAAGAAATCC 420
Db 993 CTGAGAGTGGGTGTCGAGTCTCTGTTGTTCCCTGAGCTCTGTGAAAGAAATCC 1052
QY 421 ACTTAAGTCCCTACCTGATCTCTGAGCTCTGAGCTTACCTGTTGCAACATTTGGGCAAC 480
Db 1053 ACTTAAGTCCCTACCTGATCTCTGAGCTCTGAGCTTACCTGTTGCAACATTTGGGCAAC 1112
QY 481 CGAATTCCTCCAAATCTTAATCTTGCTGCGCAGCAGATGTCCTCAACAGAGCTGATA 540
Db 1113 CGAATTCCTCCAAATCTTAATCTTGCTGCGCAGCAGATGTCCTCAACAGAGCTGATA 1172
QY 541 CAGCAGATGGGATGTTATGTTAAATGCGAGCTTACCTGTCGAAAGCTGACTTT 600
Db 1173 CAGCAGATGGGATGTTATGTTAAATGCGAGCTTACCTGTCGAAAGCTGACTTT 1232
QY 601 ATCCCGAGCTTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db 1233 ATCCCGAGCTTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1292
QY 661 CCGTGGTTGACAAATCAGTAGATTTTCATTTGAGAAAGCAAAAGCTTCAATGATGTGT 720
Db 1293 CCGTGGTTGACAAATCAGTAGATTTTCATTTGAGAAAGCAAAAGCTTCAATGATGTGT 1352
QY 721 CTAGTGCACTGTTTGTGCGGATCTCCCGCTCCGCAACATGCTATGCTTACATCATG 780
Db 1353 CTAGTGCACTGTTTGTGCGGATCTCCCGCTCCGCAACATGCTATGCTTACATCATG 1412
QY 781 AAGAGATGACATGTCTTTAGATGAGAGCTTACAGATTTGTGAAAAGAAAAGACTTACT 840
Db 1413 AAGAGATGACATGTCTTTAGATGAGAGCTTACAGATTTGTGAAAAGAAAAGACTTACT 1472
QY 841 ATATCTCCAAATCTCAATTTTCTGGGCGCACTCTGACTATGAGAAAGATTAAGAC 900
Db 1473 ATATCTCCAAATCTCAATTTTCTGGGCGCACTCTGACTATGAGAAAGATTAAGAC 1532

Query 901 CAGACT 906
Db 1533 CAGACT 1538

RESULT 13
LOCUS AX374994 3766 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 17 from Patent WO0210363.
ACCESSION AX374994
VERSION AX374994.1 GI:19169826
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Tang, Y.T., Elliott, V.S., Ramkumar, J., Yao, M.G., Burford, N., Wang, Y.E., Stewart, E.A., Gandhi, A.R., Patterson, C., Lee, E.A., Hefajia, A.J., Lu, D.A., Tribouley, C.M., Griffin, J.A., Baughn, M.R., Yue, H., Warren, B.A., Nguyen, D.B. and Walla, N.K.
TITLE Protein phosphatases
JOURNAL Patent: WO 0210363-A 17 07-FEB-2002;
Incyte Genomics, Inc. (US)
FEATURES
source location/Qualifiers
1..3766
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7480570CB1"

ORIGIN
Query Match 99.6%; Score 902.8; DB 6; Length 3766;
Best Local Similarity 99.8%; Pred. No. 5e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGGCCCATGAGTGGATGGACCAATTGTTACTGAGAGGTTGGGCTCTGCTGGAA 60
538 ATGGCCCATGAGTGGATGGACCAATTGTTACTGAGAGGTTGGGCTCTGCTGGAA 597
61 AGTGAACGGAAGAGTGTCTGCTAATTTGATAGCGGCGCATTTGGAAATCAATATCC 120
598 AGTGAACGGAAGAGTGTCTGCTAATTTGATAGCGGCGCATTTGGAAATCAATATCC 657
121 CACATTTTGAAGCCATTAAATATCACTGCTCCAGCTTATGAAGCGAGCTTGAACAG 180
658 CACATTTTGAAGCCATTAAATATCACTGCTCCAGCTTATGAAGCGAGCTTGAACAG 717
181 GACAAAGTGTATTTACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATTGAT 240
718 GACAAAGTGTATTTACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATTGAT 777
241 TGCAGTCAGAGGTTGATTTAGATCAAGCTCCCAAGATGTGGCTCTCTCTTCA 300
778 TGCAGTCAGAGGTTGATTTAGATCAAGCTCCCAAGATGTGGCTCTCTCTTCA 837
301 GACTGTTTCTCACTGTAATTTGCGGTAACCTGGAAGAAGCTTCAACTCTGTTCACTG 360
838 GACTGTTTCTCACTGTAATTTGCGGTAACCTGGAAGAAGCTTCAACTCTGTTCACTG 897
361 CTGGAGGTTGGTGTGAGTCTCTCGTGTGTTCCCTGGGCTCTGTGGAAGGAATCC 420
898 CTGGAGGTTGGTGTGAGTCTCTCGTGTGTTCCCTGGGCTCTGTGGAAGGAATCC 957
421 ACTCTAGTCCCTACCTGATTTCTCAAGCTTGTGCTTACCTGTTGCAACATTTGGCCAAC 480
958 ACTCTAGTCCCTACCTGATTTCTCAAGCTTGTGCTTACCTGTTGCAACATTTGGCCAAC 1017
481 CGAATTTTCTCCCAATCTTTATCTTGGCTGCGAGCGAGATGCTCTCAACAGAGCTGATG 540
1018 CGAATTTTCTCCCAATCTTTATCTTGGCTGCGAGCGAGATGCTCTCAACAGAGCTGATG 1077
541 CACGAGATGGAGTGGTATGTATTAATGCGCAGGATATCTGTCCAAAGCTGACTTT 600

Db 1078 CAGCAGAAATGGAGTTGGTATGTGTTAATGCCAGCAATACCTGCCAAAGCTGACTTT 1137
Query 601 ATCCCGAGTCTCATTTCTGCGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 660
Db 1138 ATCCCGAGTCTCATTTCTGCGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 1197
Query 661 CCGTGGTTGAGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCCTCAATGATGTT 720
Db 1198 CCGTGGTTGAGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCCTCAATGATGTT 1257
Query 721 CTAGTCACATGTTTAACTGCGATCTCCCGCTCCGCCACATGCTATGCGCTACATCATG 780
Db 1258 CTAGTCACATGTTTAACTGCGATCTCCCGCTCCGCCACATGCTATGCGCTACATCATG 1317
Query 781 AAGAGATGACATGCTCTTATGATGAGCTTACAGATTGTTGTGAGAAAGAAAGACTACT 840
Db 1318 AAGAGATGACATGCTCTTATGATGAGCTTACAGATTGTTGTGAGAAAGAAAGACTACT 1377
Query 841 ATATCTCCAAATCTCAATTTTCTGGGCCAACTCTGAGCATGAGAGAAATTAAAGAC 900
Db 1378 ATATCTCCAAATCTCAATTTTCTGGGCCAACTCTGAGCATGAGAGAAATTAAAGAC 1437
Query 901 CAGACT 906
Db 1438 CAGACT 1443

RESULT 14
BD171157 4790 bp DNA linear PAT 17-JAN-2003
LOCUS BD171157
DEFINITION Novel gene and protein encoded thereby.
ACCESSION BD171157
VERSION BD171157.1 GI:27876969
KEYWORDS WO 02052005-A/13.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Ohara, O., Nagase, T. and Nakajima, D.
TITLE Novel gene and protein encoded thereby
JOURNAL Patent: WO 02052005-A 13 04-JUL-2002;
KAZUSA, DNA RESEARCH INSTITUTE, OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
COMMENT OS Homo sapiens (human)
FN WO 02052005-A/13
PD 04-JUL-2002
PF 20-DEC-2001 WO 2001JP011217
PR 22-DEC-2000 JP 00P 389742
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/12,C07K14/47
CC Novel gene and protein encoded thereby
FH Key location/Qualifiers
FT CDS (184)..(2178).

FEATURES
source location/Qualifiers
1..4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.6%; Score 902.8; DB 6; Length 4790;
Best Local Similarity 99.8%; Pred. No. 5.2e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGGCCCATGAGTGGATGGACCAATTGTTACTGAGAGGTTGGGCTCTGCTGGAA 60
Db 184 ATGGCCCATGAGTGGATGGACCAATTGTTACTGAGAGGTTGGGCTCTGCTGGAA 243
Query 61 AGTGAACGGAAGAGTGTCTGCTAATTTGATAGCGGCGCATTTGGAAATCAATATCC 120
Db 244 AGTGAACGGAAGAGTGTCTGCTAATTTGATAGCGGCGCATTTGGAAATCAATATCC 303

QY 121 CACATTTTGGAGCCATTAATATCACTGCTCCAAAGCTTATGAAAGCGAAGTTGCAAG 180
 Db 304 CACATTTTGGAGCCATTAATATCACTGCTCCAAAGCTTATGAAAGCGAAGTTGCAAG 363
 QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 240
 Db 364 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 423
 QY 241 TGCAGTCAGAAAGTGTATTAACAGATCAAAAGCTCCAGATGTGCTCTCTCTTCA 300
 Db 424 TGCAGTCAGAAAGTGTATTAACAGATCAAAAGCTCCAGATGTGCTCTCTCTTCA 483
 QY 301 GACTTTTTTCTCACTGTAATCTCTGGGTAACTGGAGAAAGCTTCACTCTTCACTG 360
 Db 484 GACTTTTTTCTCACTGTAATCTCTGGGTAACTGGAGAAAGCTTCACTCTTCACTG 543
 QY 361 CTTCAGAGTGGGTTTGGCTGAGATTCTCCGCTTTTTCCTGGCTCTGTGAGAAATTC 420
 Db 544 CTTCAGAGTGGGTTTGGCTGAGATTCTCCGCTTTTTCCTGGCTCTGTGAGAAATTC 603
 QY 421 ACTCTAGTCCCTTACCTGCAATTTCTGAGCTTGTACCTGTGAGCAATTTGGCCAAAC 480
 Db 604 ACTCTAGTCCCTTACCTGCAATTTCTGAGCTTGTACCTGTGAGCAATTTGGCCAAAC 663
 QY 481 CGAATTTCTCCCAATCTTATCTGGCTGCGAGCAAGATGCTTCAACAGAGCTGATA 540
 Db 664 CGAATTTCTCCCAATCTTATCTGGCTGCGAGCAAGATGCTTCAACAGAGCTGATA 723
 QY 541 CAGCAGAAATGGGATTTGTTATGTGTTAATGCAAGCTTATCTGTGCAAAAGCTGACTT 600
 Db 724 CAGCAGAAATGGGATTTGTTATGTGTTAATGCAAGCTTATCTGTGCAAAAGCTGACTT 783
 QY 601 ATCCCGAGCTCATTTCTGCGTGTGCTGTAATGACAGCTTTTGTGAGAAATTTTG 660
 Db 784 ATCCCGAGCTCATTTCTGCGTGTGCTGTAATGACAGCTTTTGTGAGAAATTTTG 843
 QY 661 CCGTGGTTGGAACAATCGATGATTTCAATTAAGAAAGCAAAAGCTCCAAATGATGTT 720
 Db 844 CCGTGGTTGGAACAATCGATGATTTCAATTAAGAAAGCAAAAGCTCCAAATGATGTT 903
 QY 721 CTATGTCACCTTTAGCTGGGATCTCCGCTCGGCAACATGCTATGCTTACATCATG 780
 Db 904 CTATGTCACCTTTAGCTGGGATCTCCGCTCGGCAACATGCTATGCTTACATCATG 963
 QY 781 AAGAGATGACATGCTTTAGTGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
 Db 964 AAGAGATGACATGCTTTAGTGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1023
 QY 841 ATATCTTCAAACTTCAATTTCTGGGCAACTCCGAGCTATGAGAAAGATTAAGAAC 900
 Db 1024 ATATCTTCAAACTTCAATTTCTGGGCAACTCCGAGCTATGAGAAAGATTAAGAAC 1083
 QY 901 CAGACT 906
 Db 1084 CAGACT 1089

RESULT 15
 LOCUS BD183422 4790 bp DNA linear PAT 17-JUN-2003
 DEFINITION Novel genes and proteins encoded by the genes.
 ACCESSION BD183422
 VERSION BD183422.1 GI:31875622
 KEYWORDS JP 2002345492-A/135.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCES
 1 (bases 1 to 4790)
 AUTHORS Ohara,O., Nagase,T. and Nakajima,D.
 TITLE Novel genes and proteins encoded by the genes
 JOURNAL Patent: JP 2002345492-A 135 03-DEC-2002;

COMMENT KAZUSA DNA RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002345492-A/135
 PD 03-DEC-2002
 PF 26-FEB-2002 JP 200204909
 P1 OSAMU OHARA, TAKAHIRO NAGASE DAISUKE NAKAJIMA
 PC C12N15/09,C07K14/47//A61K31/711,A61K38/00,A61K48/00,A61P25/00,
 PC A61P25/14,
 PC A61P25/18,A61P35/00,C12N15/00,A61K37/02
 CC Novel genes and proteins encoded by the genes FH key
 Location/Qualifiers
 FT CDS (184).. (2178).
 FEATURES
 source Location/Qualifiers
 1..4790
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

Query Match 99.6%; Score 902.8; DB 6; Length 4790;
 Best Local Similarity 99.8%; Pred. No. 5,2e-267;
 Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
 Db 184 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
 QY 61 AGTGAACGGAAAAAGTGTGCTGTAATGATGATGATGATGATGATGATGATGATGATGAT 120
 Db 244 AGTGAACGGAAAAAGTGTGCTGTAATGATGATGATGATGATGATGATGATGATGATGAT 303
 QY 121 CACATTTTGGAAAGCCATTAATATCACTGCTCCAAAGCTTATGAGGAAAGTTGCAACAG 180
 Db 304 CACATTTTGGAAAGCCATTAATATCACTGCTCCAAAGCTTATGAGGAAAGTTGCAACAG 363
 QY 181 GACAAAGTGTATTAACAGAGCTTATGAGGAAAGCTTATGAGGAAAGTTGCAACAG 240
 Db 364 GACAAAGTGTATTAACAGAGCTTATGAGGAAAGCTTATGAGGAAAGTTGCAACAG 423
 QY 241 TGCAGTCAGAAAGTGTATTAACAGATCAAAAGCTCCAGATGTGCTCTCTCTTCA 300
 Db 424 TGCAGTCAGAAAGTGTATTAACAGATCAAAAGCTCCAGATGTGCTCTCTCTTCA 483
 QY 301 GACTTTTTTCTCACTGTAATCTCTGGGTAACTGGAGAAAGCTTCACTCTTCACTG 360
 Db 484 GACTTTTTTCTCACTGTAATCTCTGGGTAACTGGAGAAAGCTTCACTCTTCACTG 543
 QY 361 CTTCAGAGTGGGTTTGGCTGAGATTCTCCGCTTTTTCCTGGCTCTGTGAGAAATTC 420
 Db 544 CTTCAGAGTGGGTTTGGCTGAGATTCTCCGCTTTTTCCTGGCTCTGTGAGAAATTC 603
 QY 421 ACTCTAGTCCCTTACCTGCAATTTCTGAGCTTGTACCTGTGAGCAATTTGGCCAAAC 480
 Db 604 ACTCTAGTCCCTTACCTGCAATTTCTGAGCTTGTACCTGTGAGCAATTTGGCCAAAC 663
 QY 481 CGAATTTCTCCCAATCTTATCTGGCTGCGAGCAAGATGCTTCAACAGAGCTGATA 540
 Db 664 CGAATTTCTCCCAATCTTATCTGGCTGCGAGCAAGATGCTTCAACAGAGCTGATA 723
 QY 541 CAGCAGAAATGGGATTTGTTATGTGTTAATGCAAGCTTATGAGGAAAGTTGCAACAG 600
 Db 724 CAGCAGAAATGGGATTTGTTATGTGTTAATGCAAGCTTATGAGGAAAGTTGCAACAG 783
 QY 601 ATCCCGAGCTCATTTCTGCGTGTGCTGTAATGACAGCTTTTGTGAGAAATTTTG 660
 Db 784 ATCCCGAGCTCATTTCTGCGTGTGCTGTAATGACAGCTTTTGTGAGAAATTTTG 843
 QY 661 CCGTGGTTGGAACAATCGATGATTTCAATTAAGAAAGCAAAAGCTCCAAATGATGTT 720
 Db 844 CCGTGGTTGGAACAATCGATGATTTCAATTAAGAAAGCAAAAGCTCCAAATGATGTT 903
 QY 721 CTATGTCACCTTTAGCTGGGATCTCCGCTCGGCAACATGCTTACATCATG 780


```

Db      904 CTAGTGCACCTGTTAGCTGCGATCTCCCGCTCCGCCACATCGCTATCGCCTACATCATG 963
QY      781 AAGAGATGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 840
Db      964 AAGAGATGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 1023
QY      841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGGACTATGAGAAGAGATTAAAGAAC 900
Db      1024 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGGACTATGAGAAGAGATTAAAGAAC 1083
QY      901 CAGACT 906
Db      1084 CAGACT 1089

```

Search completed: June 21, 2004, 23:40:22
 Job time : 5706.52 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: June 21, 2004, 17:00:31 ; Search time 3586.84 Seconds
(without alignments)
7542.898 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 906
Sequence: 1 accgcccacagatgatgctg.....agaagattaagaccagact 906

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vtc:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rdg:*
26: em_gss_pig:*
27: em_gss_vtl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	799.8	88.3	898	13	BQ721265
2	665.2	73.4	836	12	B1821804
3	633	69.9	682	12	B1816954
4	565.8	62.5	715	14	CF721717

5	559.6	61.8	639	14	CF742387	CF742387 UT-M-HB0-
6	558.2	61.6	766	14	CF532917	CF532917 UT-M-GH0-
7	551.4	60.9	769	13	BUT04078	BUT04078 UT-M-FOO-
8	544.8	60.1	650	10	BF472046	BF472046 UT-M-BH3-
9	541.4	59.8	602	10	AM847425	AM847425 RCL-CT020
10	535.6	59.1	595	10	AM847426	AM847426 RCL-CT020
11	489.6	54.0	862	13	BQ770036	BQ770036 UT-M-F10-
12	482.8	53.3	3325	11	AK035652	AK035652 Mus muscu
13	478	47.8	792	13	BU946566	BU946566 703769910
14	430.8	47.5	946	12	BI557528	BI557528 603281865
15	428.8	47.3	512	10	BF815601	BF815601 MR2-CT012
16	422.8	46.7	516	29	AY413655	AY413655 Homo sapi
17	415.8	45.9	1060	12	BM546940	BM546940 AGENCOURT
18	404.6	44.7	619	29	CG583869	CG583869 OST226476
19	399.6	44.1	463	10	CG058779	CG058779 naflf0604
20	390.8	43.1	903	12	BG482429	BG482429 602526595
21	386.4	42.6	463	10	BF809325	BF809325 QV0-CT020
22	373.2	41.2	516	29	AY413657	AY413657 Mus muscu
23	364.2	40.2	403	10	BE004625	BE004625 CM2-BN010
24	364.2	40.2	2207	11	BC038231	BC038231 Homo sapi
25	359.4	39.7	484	13	BQ318152	BQ318152 RCL-CT046
26	359.4	39.7	494	13	BQ318155	BQ318155 RCL-CT046
27	356.4	39.3	421	10	BE346884	BE346884 UT-M-BH3-
28	352.6	38.9	1148	12	BM906608	BM906608 AGENCOURT
29	352.2	38.9	924	14	CA474739	CA474739 AGENCOURT
30	349.6	38.6	404	10	BB816652	BB816652 BB816652
31	341	37.6	466	10	AM855006	AM855006 PM3-CT026
32	337.2	37.2	442	14	CB606149	CB606149 AMGNNUC:N
33	333.8	36.8	412	9	AA879894	AA879894 vW01d07.x
34	327.6	36.2	366	10	BG002947	BG002947 RCL-CT007
35	323.2	35.7	402	29	CG499793	CG499793 OST42180
36	322	35.5	685	14	CB165612	CB165612 BE0603161
37	318	35.1	579	10	BF582366	BF582366 602094304
38	293.6	32.6	324	13	BQ367284	BQ367284 RCL-CT022
39	290	32.0	635	10	BE377720	BE377720 601229866
40	279	30.8	683	12	BM942528	BM942528 UT-M-CGDP
41	279	30.8	516	29	AY413656	AY413656 Pan trogl
42	276.6	30.5	337	14	CB606296	CB606296 AMGNNUC:S
43	248	27.4	801	14	CK130134	CK130134 AGENCOURT
44	245.6	27.1	424	10	AM461438	AM461438 B2230005A
45	239.2	26.4	881	14	CF266082	CF266082 AGENCOURT

ALIGNMENTS

RESULT 1
LOCUS BQ721265
DEFINITION AGENCOURT 8294508 lupski, sympathetic_trunk Homo sapiens cDNA clone IMAGE:6194455 5', mRNA sequence.
ACCESSION BQ721265
VERSION BQ721265.1 GI:21860162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-romail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHAM1359 row: 1 column: 08
High quality sequence stop: 669.

FEATURES

source

Location/Qualifiers
 1..898
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6194455"
 /sex="male"
 /issue_type="symptomatic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /clone_lib="Lupski: sympathetic trunk"
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site 1: NotI; Site 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors:
 5'-TGACCCAGCGGCGG-3' and
 5'-GACTAGTCTGATGCGGAGCGGCGGCTT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."

ORIGIN

Query Match
 Best Local Similarity 96.8%; Score 799.8; DB 13; Length 898;
 Matches 859; Conservative 0; Mismatches 22; Indels 6; Gaps 4;

```

OY 8 ATGAGATGATGGAATCTCAATTTGTTACTGAGAGGTTGGCTCTGCGGAAAGTGGAA 67
DB 12 ATGAGATGATGGAATCTCAATTTGTTACTGAGAGGTTGGCTCTGCGGAAAGTGGAA 71
OY 68 CGGAAAAAGTGTCTGCTAATTTAGTACCGGCAATTTGGGAATACATATCCCATTT 127
DB 72 CGGAAAAAGTGTCTGCTAATTTAGTACCGGCAATTTGGGAATACATATCCCATTT 131
OY 128 TGGAAAGCCATTATATCAAGCTCTCAAGCTTATGAAAGGAGTTGCAACAGGAAAG 187
DB 132 TGGAAAGCCATTATATCAAGCTCTCAAGCTTATGAAAGGAGTTGCAACAGGAAAG 191
OY 188 TGTAAATTAACAGAGCTATCCAGCATTCAGCGAAACATTAAGTTGATGTTGCAATG 247
DB 192 TGTAAATTAACAGAGCTATCCAGCATTCAGCGAAACATTAAGTTGATGTTGCAATG 251
OY 248 AGAAGGTTGATGATGATCAAAAGCTCCCAAGATTTGCTCTCTCTCAAGCTGTT 307
DB 252 AGAAGGTTGATGATGATCAAAAGCTCCCAAGATTTGCTCTCTCTCTCAAGCTGTT 311
OY 308 TTCTCACTGATCTTCTGGGTTAACTGAGAAAGCTTCAACTGTTGACCTGTTGAG 367
DB 312 TTCTCACTGATCTTCTGGGTTAACTGAGAAAGCTTCAACTGTTGACCTGTTGAG 371
OY 368 GTGGGTTTGTGAGTCTCTGCTGTTTCCCTGGCTCTGTTGAAAGAAATCCACTGAG 427
DB 372 GTGGGTTTGTGAGTCTCTGCTGTTTCCCTGGCTCTGTTGAAAGAAATCCACTGAG 431
OY 428 TCCCTACCTGCAATTTCTCAGCCTTGTACCTGTTGCAACATTTGGGCGCAACCCGAATTC 487
DB 432 TCCCTACCTGCAATTTCTCAGCCTTGTACCTGTTGCAACATTTGGGCGCAACCCGAATTC 491
OY 488 TTCCCAATCTTATCTTGGCTGCGAGGAGATGTCCTCAACAGAGCTGATACAGCAGA 547
DB 492 TTCCCAATCTTATCTTGGCTGCGAGGAGATGTCCTCAACAGAGCTGATACAGCAGA 551
OY 548 ATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 607
DB 552 ATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 611
OY 608 AGTCTCAATTTCTGCGCTGCTGCTGATGAGAGCTTTTGTGAAAAATTTTGGCGTGT 667
DB 612 AGTCTCAATTTCTGCGCTGCTGCTGATGAGAGCTTTTGTGAAAAATTTTGGCGTGT 670
OY 668 TGGCAAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 727

```

```

DB 671 TGGCAAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 730
OY 728 ACTGTTAGCTGGAGATCCCGCTCGGACACATGCTATGCTTACATCATGAGAGGA 787
DB 731 ACTGTTAGCTGGAGATCCCGCTCGGACACATGCTATGCTTACATCATGAGAGGA 789
OY 788 TGGCAATGTC-TTAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 846
DB 790 TGGCAATGTC-TTAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 849
OY 847 CCAAACTT--CAATTTCTGGGCAACTCTGAGCTATGAGAGGA 890
DB 850 CCAAACTTCAATTTTCTTGGGCAACTCTGAGCTATGAGAGGA 896

```

RESULT 2
 B1821804
 LOCUS
 DEFINITION
 B1821804 836 bp mRNA linear EST 04-OCT-2001
 603035883rl NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176724 5',
 mRNA sequence.
 B1821804
 VERSION
 B1821804.1 GI:15933354
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L14M11440 row: f column: 21
 High quality sequence stop: 805.

FEATURES

source

Location/Qualifiers
 1..836
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5176724"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector: pcMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, enriched for full-length clones and was constructed and Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

Query Match
 Best Local Similarity 73.4%; Score 665.2; DB 12; Length 836;
 Matches 747; Conservative 0; Mismatches 8; Indels 7; Gaps 7;

```

OY 147 CTGCTCAAGCTTATGAGAGGAGTTGCAACAGCAAAAGTTAATTAAGAGCTCAT 206
DB 1 CTGCTCAAGCTTATGAGAGGAGTTGCAACAGCAAAAGTTAATTAAGAGCTCAT 60
OY 207 CCAGCATTCAGCAAAATTAAGTTGATGATGATGATGATGATGATGATGATGATGATGATG 266
DB 61 CCAGCATTCAGCAAAATTAAGTTGATGATGATGATGATGATGATGATGATGATGATGATG 119

```

267 TCAAGCTCCAGAGATGTTGCTCTCTCTCTGAGACTGTTTCTCACTGACTTCTGGG 326
120 TCAAGCTCCAGAGATG-TGCTCTCTCTCTCTGAGACTGTTTCTCACTGACTTCTGGG 178
327 TAAAGTGAAGAGCTTCACTCTGTTCACTGCTTGAAGTGGTGGTGGTGGTGGTGGTGG 386
179 TAAAGTGAAGAGCTTCACTCTGTTCACTGCTTGAAGTGGTGGTGGTGGTGGTGGTGG 238
387 TCGTGTGTTCCCTGGCTCTGTGAAGGAAATCACTAGTCCCTGACCTGATTTCTCA 446
239 TCGTAGTTCCTGGCTCTGTGAAGGAAATCACTAGTCCCTGACCTGATTTCTCA 298
447 GCTTGTCTTACCTGTTGCTGAAGTGGGCTCAACCTGATTTCTGATTTCTGATTTCTG 506
299 GCTTGTCTTACCTGTTGCTGAAGTGGGCTCAACCTGATTTCTGATTTCTGATTTCTG 358
507 CTGCGACGAGATGCTCTCAACAGAGAGTGTATACAGACAGATGGATGGATGGATGGAT 566
359 CTGCGACGAGATGCTCTCAACAGAGAGTGTATACAGACAGATGGATGGATGGATGGAT 418
567 AATGCGACGATATACCTGCTCAAGCTTGAATTTATCCGAGTCTTCAATTTCTGCTGT 626
419 AATGCGACGATATACCTGCTCAAGCTTGAATTTATCCGAGTCTTCAATTTCTGCTGT 478
627 GCTGTGATGACAGCTTTTGTGAAGAAATTTTGGCTGGTGGACAAATCACTAGATTT 686
479 GCTGTGATGACAGCTTTTGTGAAGAAATTTTGGCTGGTGGACAAATCACTAGATTT 538
687 CATTGAGAAAGAAAGCTTCAATGATGTTTCTGATGACGATTTAGCTGGATCTG 746
539 CATTGAGAAAGAAAGCTTCAATGATGTTTCTGATGACGATTTAGCTGGATCTG 597
747 CCGCTCGGACCATGCTATGCTTCAATCATGAGAGAGATGACATGCTTTAGATGA 806
598 CCGCTCGG-CACCATGCTATGCT-CTACATCATGAGAGATGACATGCTTTAGATGA 655
807 AGCTTACAGATTTGTGAAGAA-AAAAGACTTACTATTTCCAACTTCAATTTCT-G 864
656 AGCTTACAGATTTGTGAAGAAAGAGAGCTTACTATTTCTCAAACTTCAATTTCTGG 715
865 GGGCAACTCTGAGCTATGAGAGAGATTAAGAACAGACT 906
716 GGGCAACTCTGAGCTATGAGAGAGATTAAGAACAGACT 757

RESULT 3
BI16954/c 682 bp mRNA linear EST 10-OCT-2001
LOCUS BI16954
DEFINITION imagec_10_2000/blz410bdf41.xl Soares_NPBMC Homo sapiens cDNA
clone IMAGE:4140798 3', mRNA sequence.
BI16954
BI16954.1 GI:15911639
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 682)
Kale, P.I., Harsch, T.J., Folte, P.A., Nelson, D.O., Sanders, C.G. and
Prange, C.K.
The I.M.A.G.E. Consortium quality control effort: clone
ressequencing for verification
Unpublished (2001)
Other ESTs: BG058779
Contact: Prange CK
The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA
Email: helpimage.llnl.gov
This read has been verified (found to hit its original self in the
correct orientation), as part of the I.M.A.G.E. Consortium quality
control effort. High quality sequence is defined as having 100 or

more base pairs with a phred quality value of 20 or greater, where
a sliding window of 4 base pairs with a phred quality value of 15
or greater marks the beginning and end of the sequence. For
information on obtaining this clone, please contact
info@image.llnl.gov. effort.
Plate: LLAM9388 row: k column: 7
Seq primer: -21m13
High quality sequence stop: 682.
Location/Qualifiers
1. 682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4140798"
/issue_type="lymphocyte"
/lab_host="DHI08 (phage-resistant)"
/note="Organ: blood; Vector: pT73D-Pac; Site 1: NotI;
Site 2: EcoRI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer 15'
TGTACCAATCTGAAGTGGAGCGCGCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
is normalized; constructed in the laboratory of M. Bento
Soares (University of Iowa)."

ORIGIN
Query Match 69.9%; Score 633; DB 12; Length 682;
Best Local Similarity 99.1%; Pred. No. 5.3e-178;
Matches 647; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
169 AGGTGCAACAGACAAAGTGTATTTATACAGACTATCCAGATTCAGGAAACATTAAG 228
652 AAGGTGCAACAGACAAAGTGT-ATTACAGACTATCCAGATTCAGGAAACATTAAG 594
229 GTTACATGATGATGACGAGTGAAGGTTGTATGATGATGATGATGATGATGATGATGAT 288
593 GTTACATGATGATGACGAGTGAAGGTTGTATGATGATGATGATGATGATGATGATGAT 534
289 TCTCTCTCTCAGACTGTTTCTCACTGATCTTCTGGGTTAACTGAGAAAGACTTCAAC 348
533 TCTCTCTCTCAGACTGTTTCTCACTGATCTTCTGGGTTAACTGAGAAAGACTTCAAC 474
349 TCTCTCTCTCAGACTGTTTCTCACTGATCTTCTGGGTTAACTGAGAAAGACTTCTGT 408
473 TCTCTCTCTCAGACTGTTTCTCACTGATCTTCTGGGTTAACTGAGAAAGACTTCTGT 414
409 GAGGAAATCCACTAGTCCCTACCTGATTTCCAGCTTCACTTACCTGATTTCAAC 468
413 GAGGAAATCCACTAGTCCCTACCTGATTTCCAGCTTCACTTACCTGATTTCAAC 354
469 ATTGGGCAACCGAATTTCTTCCCAATCTTATCTTGGCTGCCAGAGATGCTCAAC 528
353 ATTGGGCAACCGAATTTCTTCCCAATCTTATCTTGGCTGCCAGAGATGCTCAAC 294
529 AAGGACCTGATACAGAGAAATGGATGGTATGTTAAATGCTGATTTACTGTTCA 588
293 AAGGACCTGATACAGAGAAATGGATGGTATGTTAAATGCTGATTTACTGTTCA 234
589 AAGGACCTGATTTATCCCGAGTCTCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
233 AAGGACCTGATTTATCCCGAGTCTCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 174
649 GAGGAAATTTTGGCTGTTGGAACAATGATGATTTCAATGAGAAAGCAAGCTTCC 708
173 GAGGAAATTTTGGCTGTTGGAACAATGATGATTTCAATGAGAAAGCAAGCTTCC 114
709 AATGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 768
113 AATGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 54
769 GCTTACATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 821

Db 53 GCGTACATCATGAGAGATGACATGCTTGTAGTGAAGCTTACAGTTTGT 1

RESULT 4
CF727177 715 bp mRNA linear EST 09-OCT-2003
LOCUS CF727177
DEFINITION UI-M-HBO-ckj-j-09-0-UI.r1 NIH BMAP_HBO Mus musculus cDNA clone
IMAGE:30548096 5', mRNA sequence.
ACCESSION CF727177
VERSION CF727177.1 GI:37601345
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
DNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousef1.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES
source
Seq primer: PYX-5
Location/Qualifiers
1..715
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30548096"
/tissue="IMAG:30548096"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_HBO"
/note="Organ: Eye; Vector: PYX-Asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
to Bonaudo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to RNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into PYX-Asc vector. The library tag
is TTAGTGAAGT. This library was created for the polya tail
Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN

Query Match 62.5%; Score 565.8; DB 14; Length 715;
Best Local Similarity 92.0%; Pred. No. 7.5e-158;
Matches 597; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 258 AGTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCACTGTTTTCACAGT 317
Db 1 AGTTTATGATCAAGTTCCTCCCAAGATGTTGCTCTCTCTCTCTCTCTCTCT 317
QY 318 ACTTCTGGGTAACCTGAGAGAGAGCTTCAACTCTGTCAACCGCTTGCAGGTTGG 377
Db 61 ACTTCTGGGTAACCTGAGAGAGAGCTTCAACTCTGTCAACCGCTTGCAGGTTGG 377
QY 378 TGAATCTCTCGTGTTCCTCGGCTCTGTGAAGAAATCACTAGTCCCTACCTG 437
Db 121 TGAATCTCTCGTGTTCCTCGGCTCTGTGAAGAAATCACTAGTCCCTACCTG 180

QY 438 CATTCTCAGCCTTGTCTTACCTGTGCGCAACATGGGCGCAACCGCAATTTTCCCAATCT 497
Db 181 CATATCTCAGCCTTGTCTTACCTGTGCGCAACATGGGCGCAACCGCAATTTTCCCAATCT 240
QY 498 TTAATCTGAGCTGCGCAGCGAGATGCTCTCAACAGAGCTGTATACAGAGATGGATTTGG 557
Db 241 CTATCTTGGCTGCGCAGCGAGATGCTCTCAACAGAGCTGTATACAGAGATGGATTTGG 300
QY 558 TTAATGTTAAATGCGAGCTATACCTGTGCGCAACCGCTGAGTATCCCGAGTCTATTT 617
Db 301 CTATGTTAAATGCGAGCTATACCTGTGCGCAACCGCTGAGTATCCCGAGTCTATTT 360
QY 618 CTTGCGTGTGCTGGAATGACAGCTTTTGTGAAATTTTGGCCGTGTGGACAAATC 677
Db 361 CTTGCGGAGTGTGGAATGACAGCTTTTGTGAAATTTTGGCCGTGTGGACAAATC 420
QY 678 AGTATATTTCAATGAAAGCAAAAGCTTCAATGATGATGTTTACGCACTGTTTNGC 737
Db 421 TGTGATTTCAATGAAAGCAAAAGCTTCAATGATGATGTTTACGCACTGTTTNGC 480
QY 738 TGGATCTCCCGCTCGCGCAACATGCTATGCTTCAATGATGATGATGATGATGATGATG 797
Db 481 TGGATCTCTCGCTCGCGCAACATGCTATGCTTCAATGATGATGATGATGATGATGATG 540
QY 798 TTTAGATGAGCTTCAAGATTTTGTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 857
Db 541 TTTAGATGAGCTTCAAGATTTTGTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 600
QY 858 TTTTCTGGGCAACTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 906
Db 601 TTTTATGGGCAACTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 649

RESULT 5
CF742387 639 bp mRNA linear EST 10-OCT-2003
LOCUS CF742387
DEFINITION UI-M-HBO-cl1-c-20-0-UI.r1 NIH BMAP_HBO Mus musculus cDNA clone
IMAGE:30619363 5', mRNA sequence.
ACCESSION CF742387
VERSION CF742387.1 GI:37638726
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
DNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousef1.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

source

Location/Qualifiers
1..639
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30619363"
/tissue="IMAG:30619363"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_HBO"
/note="Organ: Eye; Vector: PYX-Asc; Site 1: Ecor I;

Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Abs vector. The library tag sequence located between the Not I site and the polyA tail is TTAATGAAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Query Match 61.8%; Score 559.6; DB 14; Length 639;
Best Local Similarity 92.4%; Pred. No. 5.1e-156;
Matches 589; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```

QY 50 CTCTGCTGGAAAGTGAACGAAAGTGTCTTAATTGATCCGCCATTGTGAAT 109
DB 1 CTCTGCTGGAAAGTGAACGAAAGTGTCTTAATTGATCCGCCATTGTGAAT 60
QY 110 ACAATACATCCCATTTTGAAGCCATTATATCACTGCTCCAACTATGAAGCAA 169
DB 61 ACAATACATCCCATTTTGAAGCCATTATATCACTGCTCCAACTATGAAGCAA 120
QY 170 GGTTCACACGACAAAGTGTATTAATACAGAGCTCATCGACATTCAGCAACATTAAG 229
DB 121 GGTTCACACGACAAAGTGTATTAATACAGAGCTCATCGACATTCAGCAACATTAAG 180
QY 230 TTACACATTTGATTCAGACGACAAAGTGTATTAATACAGAGCTCATCGACATTCAGCAAC 289
DB 181 TTACACATTTGATTCAGACGACAAAGTGTATTAATACAGAGCTCATCGACATTCAGCAAC 240
QY 290 CTCTCTCTGAGACTGTTTCTCACTGTATCTTGGGTAACTGAGAAAGCTTCACT 349
DB 241 CTCTCTCTGAGACTGTTTCTCACTGTATCTTGGGTAACTGAGAAAGCTTCACT 300
QY 350 CTGTTCACCTGCTTCGACAGTGGGTGGTGTCTGCTGTTCTGCTGCTGCTGCTG 409
DB 301 CTGTTCACCTGCTTCGACAGTGGGTGGTGTCTGCTGTTCTGCTGCTGCTGCTG 360
QY 410 AAGGAAATTCACATCTAGTCTCCATCTGATCTTCTCAAGCTTCTTACCTGTTGCCAACA 469
DB 361 AAGGAAATTCACATCTAGTCTCCATCTGATCTTCTCAAGCTTCTTACCTGTTGCCAACA 420
QY 470 TTGGGCAACCCGAAATTTCTCCCAATCTTATCTTGGCTGCAAGAGATGCTCTCAACA 529
DB 421 TTGGGCAACCCGAAATTTCTCCCAATCTTATCTTGGCTGCAAGAGATGCTCTCAACA 480
QY 530 AAGAGCTGATACAGCAAGATGGATGGTATGTATTAATGTCAGCTATATCTGTCACA 589
DB 481 AAGAGCTGATACAGCAAGATGGATGGTATGTATTAATGTCAGCTATATCTGTCACA 540
QY 590 AGCTTGAATTTATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 649
DB 541 AGCTTGAATTTATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 650 AAGAAATTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 688
DB 601 AAGAAATTCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 639

```

RESULT 6
CF532917
LOCUS
DEFINITION CF532917 656 bp mRNA linear EST 12-SEP-2003
IMAGE:3057529 5', mRNA sequence.
ACCESSION
VERSION CF532917.1 GI:34584885
KEYWORDS
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 656)
REFERENCE
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES

Seq primer: pX-5.
Location/Qualifiers
1..656
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:3057529"
/tissue_type="whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="NIH BMAP GH0"
/note="Organ: Brain; Vector: pX-Abs; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Abs vector. The library tag sequence located between the Not I site and the polyA tail is CCACTGAAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 61.6%; Score 558.2; DB 14; Length 656;
Best Local Similarity 92.4%; Pred. No. 1.4e-155;
Matches 598; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```

QY 44 TGGTGGCTCTGCTGGAAGTGAACGAAAGTGTCTTAATTGATCCGCCATT 102
DB 1 TGGTGGCTCTGCTGGAAGTGAACGAAAGTGTCTTAATTGATCCGCCATT 60
QY 103 GTGGAATACATATCCCAATTTTGAAGCATTATATCAAGCTTCCAGCTTATG 162
DB 61 GTGGAATACATATCCCAATTTTGAAGCATTATATCAAGCTTCCAGCTTATG 120
QY 163 AAGCGAAGTTGACACGACAAAGTGTATTAATTAACAGCTCATCCAGCAAG 222
DB 121 AAGCGAAGTTGACACGACAAAGTGTATTAATTAACAGCTCATCCAGCAAG 180
QY 223 CATAGCTTGAATTAATTCAGTCAAGAGTTGATTAAGATCAAGCTCCCAAG 282
DB 181 CATAGCTTGAATTAATTCAGTCAAGAGTTGATTAAGATCAAGCTCCCAAG 240
QY 283 GTTGCTCTCTCTCTGACAGTCTTCTCACTGTACTCTGGGTAACTGAGAGAGC 342
DB 241 GTTGCTCTCTCTCTGACAGTCTTCTCACTGTACTCTGGGTAACTGAGAGAGC 300
QY 343 TTCAACTCTGTTCACTGCTGACAGTGGGTGCTGAGATTCCTGTTGTTCTCTG 402
DB 301 TTCAACTCTGTTCACTGCTGACAGTGGGTGCTGAGATTCCTGTTGTTCTCTG 360

```

QY 403 CTCTGTGAAGAAAAATCCACTTACTGCTCCCTACCTGATTTCTCAGCTGCTGCTTACCTGTT 462
 Db 361 CTCTGTGAAGAAAAATCCACTTACTGCTCCCTACCTGATTTCTCAGCTGCTGCTTACCTGTT 420
 QY 463 GCCAATCTTGGGCGCAACCCGAATTTCTTCCGAATCTTTATCTTGGCTGCGCAGAGATGTC 522
 Db 421 GCGAATCTTGGGCGCAACCCGAATTTCTTCCGAATCTTTATCTTGGCTGCGCAGAGATGTC 480
 QY 523 CTCACAAAGAGCTGATTCAGACAAATGGGATTTGTTATGTTAAATGCCAGCAATACC 582
 Db 481 CTCACAAAGAGCTGATTCAGACAAATGGGATTTGTTATGTTAAATGCCAGCAATACC 540
 QY 583 TTGTCCAAAGCTGCTGATTTATCCCGAGTCTCATTTCTGCGTGGCTGTTGATGACAGC 642
 Db 541 TTGTCCAAAGCTGCTGATTTATCCCGAGTCTCATTTCTGCGTGGCTGTTGATGACAGC 600
 QY 643 TTTTGTGAGAAAAATTTTGGCGTGGTGACAAATCAGTATTTTCAT 689
 Db 601 TTTTGTGAGAAAAATCTCAGTATTTGACAAATCAGTATTTTCAT 647

RESULT 7
 BU704078 769 bp mRNA linear EST 15-JUL-2003
 LOCUS BU704078
 DEFINITION UI-M-F00-bzr-n-23-0-UI_r1 NIH_BMAP_F00 Mus musculus cDNA clone
 ACCESSION IMAGE:6406486 5', mRNA sequence.
 VERSION BU704078
 KEYWORDS EST.
 SOURCE GI:23631836
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rewmail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIN at:
 http://image.lnlnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: PYX-5.

FEATURES
 Source Location/Qualifiers
 1. 769
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6406486"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T7 phage resistant)"
 /clone_lib="NIH_BMAP_F00"
 /note="Organ: Brain; Vector: PYX-Asc; Site 1: Bcor I;
 Site 2: Not 1; The library was constructed according
 to Bontido, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 size selected according to mRNA size fraction, ligated
 with Bcor I adaptor, digested with NotI and then cloned
 directionally into PYX-Asc vector. The library tag
 is TGAAGAGCC. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

Query Match 60.9%; Score 551.4; DB 13; Length 769;
 Best Local Similarity 91.9%; Pred. No 1.6e-153;
 Matches 582; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

ORIGIN
 274 TCCCAAGATGTTGCTCTCTCTTCAAGCTTTTCTCACTGATTTCTGAGTAACTG 333
 Db 3 TTCCAAAGATGTTGCTCTCTCTGCTGACGCTTTCTCACTGATTTCTGAGTAACTG 62
 QY 334 GAGAAAGCTTCAACTCTGTTCACTGCTGAGGAGGTTTCTGAGTCTCTGTTGT 393
 Db 63 GAGAGAGCTTCAACTCTGTTCACTGCTGAGGAGGTTTCTGAGTCTCTGTTGT 122
 QY 394 TTCCCTGAGCTCTGTTGAAGAAATCACTGATGCTTACCTGATTTCTGAGTAACTG 453
 Db 123 TTCCCTGAGCTCTGTTGAAGAAATCACTGATGCTTACCTGATTTCTGAGTAACTG 182
 QY 454 TTACTCTGTTGCCAATTTGGGCCAACCAGATTTCTCCCAATCTTATCTTGGCTCCAG 513
 Db 183 TTACTCTGTTGCCAATTTGGGCCAACCAGATTTCTCCCAATCTTATCTTGGCTCCAG 242
 QY 514 CGAGATGTCCTCAACGAAGAGCTGATACGACAAATGGGATTTGTTATGTTAAATGCC 573
 Db 243 CGAGATGTCCTCAACGAAGAGCTGATACGACAAATGGGATTTGTTATGTTAAATGCC 302
 QY 574 AGCTATACCTGTTCCAAAGCTGATTTATCCCGAGTCTATTTCCGCTGCTGCTG 633
 Db 303 AGCAATACCTGTTCCAAAGCTGATTTATCCCGAGTCTATTTCCGCTGCTGCTG 362
 QY 634 AATGACAGCTTTTGGAGAAAATTTGCCGTTGGGACAAATCAGTATTTCAATTGAG 693
 Db 363 AATGACAGCTTTTGGAGAAAATTTGCCGTTGGGACAAATCAGTATTTCAATTGAG 422
 QY 694 AATGACAGCTTTTGGAGAAAATTTGCCGTTGGGACAAATCAGTATTTCAATTGAG 753
 Db 423 AATGACAGCTTTTGGAGAAAATTTGCCGTTGGGACAAATCAGTATTTCAATTGAG 482
 QY 754 GCCAATCTGCTATGCTCATCATCATGAGATGAGATGATCTTTAGATGAGCTTAC 813
 Db 483 GCCAATCTGCTATGCTCATCATCATGAGATGAGATGATCTTTAGATGAGCTTAC 542
 QY 814 AGATTGTGAAAAGAAAAAGACTTACTATATCTCCAACTTCAATTTTCTGGGCACTC 873
 Db 543 AGATTGTGAAAAGAAAAAGACTTACTATATCTCCAACTTCAATTTTCTGGGCACTC 602
 QY 874 CTGACTATGAGAAAGAAATTAAGAACCACT 906
 Db 603 ATGACTATGAGAAAGAAATTAAGAACCACT 635

RESULT 8
 BF472046 650 bp mRNA linear EST 04-DEC-2000
 LOCUS BF472046
 DEFINITION UI-M-BH3-awu-d-10-0-UI_r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
 ACCESSION UI-M-BH3-awu-d-10-0-UI_5', mRNA sequence.
 VERSION BF472046
 KEYWORDS EST.
 SOURCE GI:11541229
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rewmail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIN at:
 http://image.lnlnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: PYX-5.

COMMENT

Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: MEST@mail.nih.gov
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
 should be noted that Bento Soares is generating a small number of
 additional specialized non-redundant arrays of BMAP cDNAs whose
 availability will be considered under appropriate and limited
 collaborative arrangements
 Seq primer: M13 Reverse

FEATURES

Source

Location/Qualifiers
 1..650
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-aw-d-10-0-UI"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP M S4"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The
 NIH BMAP M S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH BMAP M S4,
 NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S4,
 NIH BMAP M S2, NIH BMAP M S1. The subtracted library
 (NIH BMAP M S4) was constructed as follows: PCR amplified
 cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2, and
 NIH BMAP M S3.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH BMAP M S4 library. This procedure has been previously
 described (Bonaldi, Lennon and Soares, Genome Research
 6:791-806, 1996)"

ORIGIN

Query Match 60.1%; Score 544.8; DB 10; Length 650;
 Best Local Similarity 93.1%; Pred. No. 1,4e-151;
 Matches 603; Conservative 0; Mismatches 42; Indels 3; Gaps 3;

11 AGATGATGGAACTCAATTTGATCTGAGAGGTTGTGCTCTGCGAAGTGAACGG 70
 1 AGATGATGGAACTCAATTTGATCTGAGAGGTTGTGCTCTGCGAAGTGAACGG 60
 71 AAAAGGCTGCTAATGATGAGCGGCAATTTGGAATACATATCCCAATTTGG 130
 61 AAAAA-TGCGCTAATGATGAGCGGCAATTTGGAATACATATCCCAATTTGG 119
 131 AAGCCATTAATCACTGCTCAAGCTTATGAAGCGAAGTTGCAACAGAAAGTGT 190
 120 AAGCCATTAATCACTGCTCAAGCTTATGAAGCGAAGTTGCAACAGAAAGTGT 179
 191 TAATTACAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGATGAGTCAGA 250
 180 TAATTACAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGATGAGTCAGA 239
 251 AGGTGATGTTAGATCAAGCTCCCAAGATGTGCTCTCTCTTCAGACTGTTTTC 310

Db 240 AGGTGATGTTAGATCAAGCTCCCAAGATGTGCTCTCTCTTCAGACTGTTTTC 299
 Qy 311 TCACGTACTTCTGGGTAAACGTGAGAAAGCTTCAACTCTGTTCACTCTGCAGCTG 370
 Db 300 TCACGTACTTCTGGGTAAACGTGAGAAAGCTTCAACTCTGTTCACTCTGCAGCTG 359
 Qy 371 GGTTCGTAGTCTCTCTGTTTTCCTGCGGCTCTGTGAAAGAAATCACTCTAGTCC 430
 Db 360 GCTTCGTAGTCTCTCTGTTTTCCTGCGGCTCTGTGAAAG-AAAATCACTCTAGTCC 418
 Qy 431 CTACCTGATTTCTCAAGCTTGTGCTTACCTGTTGCCAATTTGGGCCAACCGAATCTTC 490
 Db 419 CTACCTGATTTCTCAAGCTTGTGCTTACCTGTTGCCAATTTGGGCCAACCGAATCTTC 478
 Qy 491 CCAATCTTATCTTGGCTGCGACGAGATGCTCTCAACAGAGCTGATACAGCAGATG 550
 Db 479 CCAATCTTATCTTGGCTGCGACGAGATGCTCTCAACAGAGCTGATACAGCAGATG 538
 Qy 551 GCAATGCTTATGTTTAAATGCAAGCTATACCTGTCGAAGCCTGACTTATCCGAGT 610
 Db 539 GCAATGCTTATGTTTAAATGCAAGCTATACCTGTCGAAGCCTGACTTATCCGAGT 598
 Qy 611 CTCAATTCCTGCGCTGCTGCTGTAATGACAGCTTTTGTGAGAAATTT 658
 Db 599 CTCAATTCCTGCGAGTCTGTAATGACAGCTTTTGTGAGAAATTT 645

RESULT 9
 AM847425/c
 LOCUS 602 bp mRNA linear EST 19-MAY-2000
 DEFINITION R1-CT0206-270999-021-g01 CT0206 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM847425
 VERSION AM847425.1 GI:7942942
 KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 602)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=ct2-R1-CT0206-270

999-021-g01&ct3=1999-09-27&cl=1)
 Seq primer: puc 18 forward
 High quality sequence start: 8
 High quality sequence stop: 602.

Location/Qualifiers

1..602

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="CT0206"

/note="Organ: Colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Query Match 59.8%; Score 541.4; DB 10; Length 602;
Best Local Similarity 98.0%; Pred. No. 1.4e-150;
Matches 580; Conservative 0; Mismatches 7; Indels 5; Gaps 3;

169 AGCTTGCAACGACGAAAGTGTAAATACAGAGCTATCCAGCATTCAGCAAAATAG 228
195 AGCGTGCAACGACGAAAGTGTAAATACAGAGCTATCCAGCATTCAGCAAAATAG 536
229 GTTACATTTGATTCAGTCAAGAGTGTATGTTACGATCAAGCTCCCAAGATGTC 288
535 GTTACATTTGATTCAGTCAAGAGTGTATGTTACGATCAAGCTCCCAAGATGTC 477
289 TCTCTCTTCAACGCTGTTTCTCACTGATCTTGGGTAACCTGGAAGAGCTTCAAC 348
476 -CTCTCTTCAACGCTGTTTCTCACTGATCTTGGGTAACCTGGAAGAGCTTCAAC 418
349 TCTGTTACCTGCTGACAGTGGGTTTGCTGATGTTCTGTTGTTTCCCTGGCTCTGT 408
417 TCTGTTACCTGCTGACAGTGGGTTTGCTGATGTTCTGTTGTTTCCCTGGCTCTGT 358
409 GAAGAAATTCACCTTCTAGTCCCTACCTGATTTCTGAGCTTCTTACCTGTTGGCAAC 468
357 GAAGAAATTCACCTTCTAGTCCCTACCTGATTTCTGAGCTTCTTACCTGTTGGCAAC 298
469 ATTGGGCCAACCCGAAATCTTCCCAATCTTATCTGGCTCCAGCGAAGTCTCTCAAC 528
297 ATTGGGCCAACCCGAAATCTTCCCAATCTTATCTGGCTCCAGCGAAGTCTCTCAAC 238
529 AAGAGCTGATACGCAAGAAATGGATGTTATGTTAAATCCAGCTATACCTGTCAAC 588
237 AAGAGCTGATACGCAAGAAATGGATGTTATGTTAAATCCAGCTATACCTGTCAAC 178
589 AAGCTGACCTTATCCCGAGTCTCATTTCTGCGGTGCTGTTGAATGACAGCTTTTGT 648
177 AAGCTGACCTTATCCCGAGTCTCATTTCTGCGGTGCTGTTGAATGACAGCTTTTGT 118
649 GAGAAATTTTGGCGGTGTTGCAATCAATGATTTGATTTGAGAAAGAAAGCTTCC 708
117 GAGAAATTTTGGCGGTGTTGCAATCAATGATTTGATTTGAGAAAGAAAGCTTCC 708
709 AATGATGTTTGTAGTGAATGATTTGATTTGAGTGGGATCTCCGCTCCGACCA 760
60 CATGATGTTTGTAGTGAATGATTTGATTTGAGTGGGATCTCCGCTCCGACCA 9

RESULT 10
AM847426/c 595 bp mRNA linear EST 19-MAY-2000
LOCUS R1-CT0206-270999-021-906 CT0206 Homo sapiens cDNA, mRNA sequence.
DEFINITION AM847426
ACCESSION AM847426.1 GI:7942943
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,
Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL
MEDLINE
PUBMED
COMMENT
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=kt2=RC1-CT0206-270
999-021-906&cl3=1999-09-27&cl4=1)
Seq primer: puc 18 forward
High quality sequence stop: 595.

FEATURES

source
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0206"
/note="Organ: Colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 59.1%; Score 535.6; DB 10; Length 595;
Best Local Similarity 97.6%; Pred. No. 7.6e-149;
Matches 576; Conservative 0; Mismatches 9; Indels 5; Gaps 3;

169 AGCTTGCAACGACGAAAGTGTAAATACAGAGCTATCCAGCATTCAGCAAAATAG 228
585 AGCGTGCAACGACGAAAGTGTAAATACAGAGCTATCCAGCATTCAGCAAAATAG 526
229 GTTACATTTGATTCAGTCAAGAGTGTATGTTACGATCAAGCTCCCAAGATGTC 288
525 GTTACATTTGATTCAGTCAAGAGTGTATGTTACGATCAAGCTCCCAAGATGTC 467
289 TCTCTCTTCAACGCTGTTTCTCACTGATCTTGGGTAACCTGGAAGAGCTTCAAC 348
466 -CTCTCTTCAACGCTGTTTCTCACTGATCTTGGGTAACCTGGAAGAGCTTCAAC 408
349 TCTGTTACCTGCTGACAGTGGGTTTGCTGATGTTCTGTTGTTTCCCTGGCTCTGT 408
407 TCTGTTACCTGCTGACAGTGGGTTTGCTGATGTTCTGTTGTTTCCCTGGCTCTGT 348
409 GAAGAAATTCACCTTCTAGTCCCTACCTGATTTCTGAGCTTCTTACCTGTTGGCAAC 468
347 GAAGAAATTCACCTTCTAGTCCCTACCTGATTTCTGAGCTTCTTACCTGTTGGCAAC 288
469 ATTGGGCCAACCCGAAATCTTCCCAATCTTATCTGGCTCCAGCGAAGTCTCTCAAC 528
287 ATTGGGCCAACCCGAAATCTTCCCAATCTTATCTGGCTCCAGCGAAGTCTCTCAAC 228
529 AAGAGCTGATACGCAAGAAATGGATGTTATGTTAAATCCAGCTATACCTGTCAAC 588
227 AAGAGCTGATACGCAAGAAATGGATGTTATGTTAAATCCAGCTATACCTGTCAAC 168
589 AAGCTGACCTTATCCCGAGTCTCATTTCTGCGGTGCTGTTGAATGACAGCTTTTGT 648
167 AAGCTGACCTTATCCCGAGTCTCATTTCTGCGGTGCTGTTGAATGACAGCTTTTGT 108
649 GAGAAATTTTGGCGGTGTTGCAATCAATGATTTGATTTGAGAAAGAAAGCTTCC 708
107 GAGAAATTTTGGCGGTGTTGCAATCAATGATTTGATTTGAGAAAGAAAGCTTCC 51

Qy 709 AATGAGTGTCTAGTSCACTGTTAGTCGGATCTCCGCTCCGAC 758
Db 50 CATGATGTGTCTAGTSCACTGTTAGTCGGATCTCCGCTCCGAC 1

RESULT 11
B0770036 862 bp mRNA linear EST 26-JUL-2002
LOCUS B0770036
DEFINITION U1-M-F10-byt-o-24-0-U1 r1 NIH_BMAP_F10 Mus musculus cDNA clone
IMAGE:5702255 5', mRNA sequence.
ACCESSION B0770036
VERSION B0770036.1 GI:21978510
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 862)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rgs@bgl-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/MLMW at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pyX-5.
Location/Qualifiers
1..862
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5702255"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_id="NIH_BMAP_F10"
/note="Organ: Brain; Vector: pyX-Aac; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pyX-Aac vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCCAGCAGC. This library was created for the University
Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN
Query Match 54.0%; Score 489.6; DB 13; Length 862;
Best Local Similarity 92.1%; Pred. No. 5.6e-135;
Matches 516; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 347 ACTGCTTACCTGTTGAGTGGGTTGCTGATTCCTCGTTTCCCTGACCT 406
Db 9 ACTGCTTACCTGTTGAGTGGGTTGCTGATTCCTCGTTTCCCTGACCT 68

Qy 407 GTGAAGAAATTCACCTTACCTCCTACCTGACATTTCTACGCTTGTGCA 466
Db 69 GTGAAGAAATTCACCTTACCTCCTACCTGACATTTCTACGCTTGTGCA 128

Qy 467 AATTGGGCGCAACCCGAATCTTCCCAATCTTATCTTGCTGCGAGAGATGCTCA 526
Db 129 AATTGGGCGCAACCCGAATCTTCCCAATCTTATCTTGCTGCGAGAGATGCTCA 188

Qy 527 ACAAGAGCTGATACACAGAAATGGATGGTATGTTAAATCCAGCTATACCTGTC 586
Db 189 ACAAGAGCTGATACACAGAAATGGATGGTATGTTAAATCCAGCTATACCTGTC 248

Qy 587 CAAGCTGACTTATCTCCGAGTCTCATTTCTGCTGCTGCTGTGAATGACGCTTTT 646
Db 249 CAAGCTGACTTATCTCCGAGTCTCATTTCTGCTGCTGCTGTGAATGACGCTTTT 308

Qy 647 GTGAGAAATTTGCGCTGGTGGGCAAAATCAGTATTCATTGAGAAAGCAAAAGCT 706
Db 309 GTGAGAAATTTGCGCTGGTGGGCAAAATCAGTATTCATTGAGAAAGCAAAAGCT 368

Qy 707 CCAATGATGTGTTCTATGTCACATGTTTACCTGGATCTCCGCTCCGCACTATGCTA 766
Db 369 CCAATGATGTGTTCTATGTCACATGTTTACCTGGATCTCCGCTCCGCACTATGCTA 428

Qy 767 TCGCTCATCATGAGAGAGATGACATGCTTTAATGAAGCTTACAGATTGTGAAG 826
Db 429 TCGCTCATCATGAGAGAGATGACATGCTTTAATGAAGCTTACAGATTGTGAAG 488

Qy 827 AAAAAAGCTACTATATCTCCAAACTTCAATTTCTGGGCGCAACTCTGACTATGAGA 886
Db 489 AAAAAAGCTACTATATCTCCAAACTTCAATTTCTGGGCGCAACTCTGACTATGAGA 548

Qy 887 AGAAGATTAAACACAGACT 906
Db 549 AGAAGATTAAACACAGACT 568

RESULT 12
AK035652 3325 bp mRNA linear HTC 19-SEP-2003
LOCUS AK035652
DEFINITION Mus musculus adult male urinary bladder cDNA, RIKEN full-length
enriched library, clone:9530081f05 product:TRNCRATED MARK
PHOSPHATASE 7 homolog [Homo sapiens], full insert sequence.
ACCESSION AK035652
VERSION AK035652.1 GI:26330815
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayashizaki, Y., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subfraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishii, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Harada, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Maruyama, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel capillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913

PUBMED 11076861
4
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3325)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, I., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, Fax: 81-45-503-9216)

COMMENT
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 3325
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:9530081F05"
/db_xref="MGI:2399973"
/db_xref="taxon:10090"
/clone="9530081F05"
/sex="male"
/tissue_type="urinary bladder"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
460..1080
/note="unnamed protein product; 'PUNCATED MAPK PHOSPHATASE 7 homolog (Homo sapiens) (SPTK196Q82, evidence: FASTY, putative"
/codon_start=1
/protein_id="BAC29138.1"
/db_xref="GI:26330816"
/translation="MAHEMIGTQIVTESLVALNLSGNEKVLIDSRPEVYNTHTILE AININCSKLMKRLQODKVLITELIQSAAKVIDNOKVYVYDDSDVSSLSDC FLVTLGLERSFNSVHLAAGFAEPSPGRCERKSTLVPFCISQDCLPVPANIGPT RILPNLYICQDVLNKKVASKSLSLCTSVASGKAHVYHLSWNTL"

ORIGIN

Query Match 53.3%; Score 482.8; DB 11; Length 3325;
Best Local Similarity 94.0%; Pred. No. 1.3e-132;
Matches 502; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
1 ATGCCCATGATGATGGAACTCAATGTGTACTGAGAGGTGGCTGCTGCTGGA 60

|||||
460 ATGCCCATGATGATGGAATCTCAATGTTGTTACTGAGAGCTGGTGGCTCTGGA 519
61 AGTGGAAACGGAATAATGCTGCTAAATGATAGCCGCGCATTTGTGAATAACATACATCC 120
520 AGTGAACGGAATAATGCTGCTAAATGATAGCCGCGCATTTGTGAATAACATACATCC 579
121 CACATTTTGAAGCCATTAAATATCAATGCTTCAAGCTTAAAGCCGAAGTTGCAACAG 180
580 CACATTTTGAAGCCATTAAATATCAATGCTTCAAGCTTAAAGCCGAAGTTGCAACAG 639
181 GACAAAGTTTAAATACAGAGCTATCCAGCATTCAGCGAAATAGTTGACATTTGAT 240
640 GACAAAGTTTAAATACAGAGCTATCCAGCATTCAGCGAAATAGTTGACATTTGAT 699
241 TCCAGTCAGAAAGTTGATAGTTTAAATACAGAGCTATCCAGCATTCAGCGAAATAG 300
700 TCCAGTCAGAAAGTTGATAGTTTAAATACAGAGCTATCCAGCATTCAGCGAAATAG 759
301 GACTGTTTCTCACTGTTACTTCTGGTAAATCTGGAAGAGCTTCACTGTTCACTG 360
760 GACTGTTTCTCACTGTTACTTCTGGTAAATCTGGAAGAGCTTCACTGTTCACTG 819
361 CTTCAGAGTGGTGTGCTGAGTTCTCTGTTGTTCCCTGGGCTCTGGAAGAAATTC 420
820 CTTCAGAGTGGTGTGCTGAGTTCTCTGTTGTTCCCTGGGCTCTGGAAGAAATTC 879
421 ACTTAGTCCCTACCTGCAATTTCTCAGCTTCTTACCTGTTGCAATTTGGGCAAC 480
880 ACTTAGTCCCTACCTGCAATTTCTCAGCTTCTTACCTGTTGCAATTTGGGCAAC 939
481 CGAATTCCTCCAAATCTTATCTGCTGCGCAGCAGAGATGTCCTCAACAGAG 534
940 CGAATTCCTCCAAATCTTATCTGCTGCGCAGCAGAGATGTCCTCAACAGAG 993

RESULT 13
BU946569/c 792 bp mRNA linear EST 17-OCT-2003
LOCUS 70376991.001 RALUTX1.01 Rattus norvegicus cDNA, mRNA sequence.
DEFINITION BU946569
ACCESSION BU946569
VERSION BU946569.1 GI:37701866
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 792)
Fu, G.K. and Stuve, L.L.
An improved method for the construction of cDNA libraries for highly efficient DNA sequencing from the 3' end of expressed genes
Unpublished (2003)
JOURNAL Contact: Fu GK
COMMENT Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel.: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1. 792
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="lung"
/clone_lib="RALUTX1.01"
/note="Rat, lung, Clofibrate, 250 mg/kg, 12hr-day 28 sacrifice, SD, M/F, Pool"

FEATURES
source
1. 792
Location/Qualifiers

ORIGIN

Query Match 47.8%; Score 433; DB 13; Length 792;
Best Local Similarity 86.2%; Pred. No. 4.9e-118;
Matches 525; Conservative 0; Mismatches 80; Indels 4; Gaps 4;

Oy	242	GCAGCAGAAAGGTGTAGTTTACATGATCAAAAGCTCCCAATATGTGC - CTCCTCTCTTCA	300
Db	781	GCAGTCAGAAAGGTGTAGTGTATGATCAAGATTTCCAGATGTGGTTCTTCGGTCTGACAA	722
Oy	301	GACTGTTTTCTCAGTGTACTTCGTGGGTAAACTGAGAAAGCTTCAACTCT - GTTCACCT	359
Db	721	CCGGCTTCTTCACTGTGTCTTCGGGTAGCTGAGAAAGAGCTTCAACTCTGGTTCACT	662
Oy	360	GCTTGACAGTGGGTTTGC - TGAAGTTCTCTCTGTGTGTTTCCCTGGCCTCTGTGAAGAAAT	418
Db	661	GCTTGACAGTGGGTTTGTGTAGTTCTCTGTGTGTTTCCCTGGCCTCTGTGAAGAAAT	602
Oy	419	CCACTGTAGTCCCTACCTGCATTTCTCAGCCTTGCTTACCTGTGCAACATTGGGCCAA	478
Db	601	CCACTGTAGTCCCTACCTGCATATCTCAGCCTTGCTTACCTGTGCAACACTGGGCCAA	542
Oy	479	CCCGAATTTCTCCCAATCTTTATCTTGCGCTGCACGAGATGTCTCTCAACAGAGCTGA	538
Db	541	CTCGAATTTCTCCCAATCTCTATCTTCGCGCTGCAGAGATGTCTCTCAACAGAGCTGA	482
Oy	539	TACAGCAGAAATGGGATTTGGTTATGTGTAAATGCCAGCTATCTCTGCCAAAGCCTGACT	598
Db	481	TGCACAGAAACGGGATTTGGCTATGTGTCTAAATGCCAGCAACCTGTCCAAAGCCTGACT	422
Oy	599	TTATCCCGAGTCTCATTTCTCGCGTGTGCCTGTGAATGACAGCTTTGTGAAAAATTT	658
Db	421	TCATATCCGGAATCTCACTTCTCGCGAGTGCCTGTGAATGACAGCTTTGTGAAAAATCC	362
Oy	659	TGCGGTGTGTGACAAATCAGTAGAATTTCAATTGAGAAAGCAAAAGCCTTCAATGTGATGTG	718
Db	361	TGCC - TGGTTGACAAGTCTGTGGAATTTCAATTGAGAAAGCAAAAGCCTTCAACGCGCTGTG	303
Oy	719	TTCTTAGTGCATGTTTAGCTGGGATCTCCGCGTCCGCCACATCGCTATCGGCTACATGA	778
Db	302	TGCTGATTCACGTGCTTTGGCTGGGATCTCCCGCTCGGCCACCAATGTCTATTCCTATATGA	243
Oy	779	TGAAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAAGAAAAAGACTTA	838
Db	242	TGAAGAGATGACATGTCTCTGATGATGAGGCTTACAGTAAAGGTAGCTCAGAAAGGCTTA	183
Oy	839	CTATATCTC 847	
Db	182	CTCCTTCCC 174	

RESULT	14
LOCUS	B1657528
DEFINITION	B1657528 946 bp mRNA linear EST 12-SEP-2007 603281818666r1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5326049 5'
ACCESSION	B1657528
VERSION	B1657528.1 GI:15571764
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 946) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
AUTHORS	Contact: Robert Straubeberg, Ph.D.
TITLE	Email: cgapbs-remail.nih.gov
JOURNAL	Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch Ph.D.
COMMENT	cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLML1827 row: d column: 18 High quality sequence stop: 704.

FEATURES	SOURCE	Location/Qualifiers
		1. 946
		/organism="Mus musculus"
		/mol_type="mRNA"
		/strain="NMRI"
		/db_xref="taxon:10090"
		/clone="IMAGE:5326049"
		/tissue_type="tumor, gross tissue"
		/dev_stage="5 months"
		/lab_host="DH10B"
		/clone_id="NCI CGAP Mam4"
		/note="Organ: mammary; Vector: PCMV-SpORe6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furch, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."
ORIGIN		
Query Match	47.5%	Score 430.8; DB 12; Length 946;
Best Local Similarity	91.6%	Ident. No. 2.5e-117;
Matches 467; Conservative 0;	Mismatches 42;	Indels 1; Gaps 1
QY	398	CTGGCGCTTGGAAGAAATCCACTAGACCTGACCTGATTTCTGAGCTTGCTTAC
DB	1	CTGGCGCTTGGAAGAAATCCACTAGACCTGACCTGACCTGACCTGACCTGCTTAC
QY	458	CTGTTGGCCAACTTTGGGCGCAACCCGAATCTTCCCAATCTTATCTTGAGCTGAC
DB	61	CTGTTGGCAACATTTGGGCGCAACCTGAAATCTTCCCAATCTGATCTTGAGCTGAC
QY	517	GATGCTCTCAACAGAGCTGATACAGCAAGTGGATGGTATGTTGTTAAATGCCAGC
DB	121	GATGCTCTCAACAGAGCTGATACAGCAAGTGGATGGTATGTTAAATGCCAGC
QY	577	TATACCTGTCAAAGCTGACTTTATCCCGAGTCTATTTCCGCGTGGCTGCTGAT
DB	181	AATACCTGTCAAAGCTGACTTTATCCCGAGTCTATTTCCCGAGTGGCTGCTGAT
QY	637	GACAGCTTTTGGAAGAAATTTTGGCCGTGGTGGCAAAATCAGTATGTTCAATGAGAA
DB	241	GACAGCTTTTGGAAGAAATTTTGGCCGTGGTGGCAAAATCAGTATGTTCAATGAGAA
QY	697	GCAAAAGCTTCAATGATGTTGTTCTAGTCACTGTTTAACTGGGATCTCCGCTCGCC
DB	301	GCAAAAGCTTCAATGATGTTGTTCTAGTCACTGTTTAACTGGGATCTCTGCTCGCC
QY	757	ACCAATGCTATGCTTATCATATGAAAGAGTGAACATGCTTATGATGAAGCTTACGA
DB	361	ACTATGCTATGCTTATCATATGAAAGAGTGAACATGCTTATGATGAAGCTTACGA
QY	817	TTTGTGAAAAGAAAAGAGCTACTATATCTCCAAATCTCAATTTTCTGGGCCACTCTG
DB	421	TTTGTGAAAAGAAAAGAGCTACTATATCTCCAAATCTCAATTTTCTGGGCCACTCTG
QY	877	GACTATGAGAAAGATTATAGAACCAACT 906
DB	481	GACTATGAGAAAGATTATATACCAACT 510
RESULT 15		
BF815601/c		
LOCUS	BF815601	512 bp mRNA linear EST 13-JAN-2001
DEFINITION	MR2-CT0128-041200-009-c03 CT0128 Homo sapiens cDNA, mRNA sequence.	
ACCESSION	BF815601	
VERSION	BF815601.1	GI:12150300
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 512)	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
AUTHORS		

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,F., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brenan,R.K., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800

Contact: Simpson A.J.C.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-27049922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&cl=MR2-C10128-
041200-009-c03&cl3=2000-12-04&cl4=1)
Seq primer: puc 18 forward
High quality sequence start: 20
High quality sequence stop: 233.
Location/Qualifiers

FEATURES
source

1. .512
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="C10128"
/note="Organ: colon_ins; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 47.3%; Score 428.8; DB 10; Length 512;
Best Local Similarity 94.9%; Pred. No. 76-117;
Matches 486; Conservative 0; Mismatches 22; Indels 4; Gaps 4;

QY 201 GCTCATCCGCACTTACGCGAAACATPAAGTTGACATTGACATGACATGACGAAAGTTGTAGT 260
DB 512 GCTCATCCGCACTTACGCGAAACATPAAGTTGACATTGACATGACATGACGAAAGTTGTAGT 453
QY 261 TTACGATCAAGCTCCCAAGATGTGGCTCTCTCTCAGACGTCTTTCACCTGACT 320
DB 452 TTACGATCAAGCTCCCAAGATGTGGCTCTCTCTCAGACGTCTTTCACCTGACT 393
QY 321 TCTGGGTAACTGAGAAAGAGCTTCACTCTGTTCACTGCTTGACAGGTGGGTTTGCTGA 380
DB 392 TCTGGGTAACTGAGAAAGAGCTTCACTCTGTTCACTGCTTGACAGGTGGGTTTGCTGA 333
QY 381 GTTCTC-TGGTTGTTTCCCTGGCCCTCTGTGAAGGAAATTCACCTCTAGTCCCTACCTGCA 439
DB 332 GTTCTCAGTGTGTCTCTGGCTCTGTGAAGGAAATTCACCTCTAGTCCCTACCTGCA 273
QY 440 TTTCTACGCTTGTCTTACCTGTGGCAACATTTGGGCCAACCCGAATCTTCCCAATCTTT 499
DB 272 TTTCTACGCTTGTCTTACCTGTGGCAACATTTGGGCCAACCCGAATCTTCCCAATCTTT 213
QY 500 ATCTGGCTGCGCAGGAGAG-TGTCTCAACAAGAGCTGATACAGCA-GAATGGGATTGG 557
DB 212 ATCTGGCTGCGCAGGAGAGCTGCTCTCAACAAGAGCTGATGACAGACGAATGGGATTGG 153
QY 558 TTATGTGTTAAATGACGATATACCTGTCCAAAGCTGACTTATCCACGAGTTCATT 616
DB 152 TTATGTGTTAAATGACGATATACCTGTCCAAAGCTGACTTATCCACGAGTTCATT 93

QY 617 TCTGCGTGTGCTCTGTGAATGACAGCTTTGTGAGAAATTTTGCCTGTGTGACAAAT 676
DB 92 TCTGCGTGTGCTCTGTGAATGACAGCTTTGTGAGAAATTTTGCCTGTGTGACAAAT 33
QY 677 CAGTAGATTGATTGAGAAAGCAAAAGCCTCC 708
DB 32 CAGTAGATTGATTGAGAAAGCTTAAGCCTTC 1

Search completed: June 22, 2004, 02:51:56
Job time : 3588.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 12:17:10 ; Search time 526.548 Seconds
(without alignments)
7309.622 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 906

Sequence: 1 atggcccatgagatgatttg.....agaagattaaagaccagact 906

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

N_Geneseq_29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	906	100.0	5450	6	ACC60559 Polynucle
2	906	100.0	5450	6	ACC60572 Polynucle
3	902.8	99.6	2118	4	AAEF30479 Human pro
4	902.8	99.6	2732	4	AAAD09492 Human SGP
5	902.8	99.6	3059	6	AA515768 CDNA sequ
6	902.8	99.6	3496	6	ABK47596 CDNA enco
7	902.8	99.6	3544	5	AA514639 Human CDN
8	902.8	99.6	3544	6	ABK49402 CDNA enco
9	902.8	99.6	3766	6	ABK14474 Human pro
10	902.8	99.6	4790	6	ABN83966 Human gen
11	902.8	99.6	5145	5	ABV20833 Human pro
12	902.8	99.6	5145	5	ABV21080 Human pro
13	902.8	99.6	5145	5	ABV26680 Human pro
14	902.8	99.6	5145	5	ABV20978 Human pro
15	902.8	99.6	5145	5	ABV21092 Human pro
16	902.8	99.6	5145	5	ABV21312 Human pro
17	902.8	99.6	5145	5	ABV21316 Human pro
18	902.8	99.6	5145	5	ABV26826 Human pro
19	902.8	99.6	5145	5	ABV27131 Human pro
20	902.8	99.6	5145	5	ABV26923 Human pro
21	902.8	99.6	5145	5	ABV27135 Human pro
22	902.8	99.6	5145	5	ABV28657 Human pro
23	902.8	99.6	5145	5	ABV22827 Human pro

24	902.8	99.6	5145	5	ABV26934 Human pro
25	901.2	99.5	2102	7	ADA53105 Human cod
26	901.2	99.5	3104	6	ABN59704 Novel hum
27	901.2	99.5	5111	6	ACC60521 Polynucle
28	899.6	99.3	2966	4	AAH99685 Human pro
29	782.8	86.4	2756	6	ACC60560 Polynucle
30	727.6	80.3	749	4	AAH06539 Human CDN
31	564.8	62.3	3332	6	ABK48378 CDNA enco
32	411.2	45.4	418	5	ABV11252 Human pro
33	410.2	45.3	461	5	ABV40981 Human pro
34	410.2	45.3	461	5	ABV32398 Human pro
35	410.2	45.3	461	5	ABV41327 Human pro
36	395.4	43.6	427	5	ABV10887 Human pro
37	365.8	40.4	2377	7	ACA64956 Human pro
38	365.8	40.4	2377	7	ABX10760 Human pro
39	363.6	40.1	2453	2	AAH6758 CDNA of t
40	363.6	40.1	2453	6	AAH41236 Murine ne
41	359.8	39.7	419	5	ABV32042 Human pro
42	351	38.7	2415	2	AAH6757 CDNA of t
43	349	38.5	422	5	ABV02083 Human pro
44	340.8	37.6	346	5	ABV11245 Human pro
45	339.8	37.5	345	5	ABV02076 Human pro

ALIGNMENTS

RESULT 1	ACCE60559	ACCE60559 standard; CDNA; 5450 BP.
ID	ACCE60559	
AC	ACCE60559	
XX		
DT	19-JUN-2003	(first entry)
XX		
DE	Polynucleotide relating to the invention SEQ ID NO: 108.	
XX		
KW	Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;	
KW	antiproliferative; cardiatic; cytoskeletal; gene therapy; liver disease;	
KW	proliferative disorder; renal failure; cardiovascular disorder;	
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;	
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200257460-A2.	
XX		
PD	25-JUL-2002.	
XX		
PF	20-DEC-2001; 2001WO-US050459.	
XX		
PR	20-DEC-2000; 2000US-0256686P.	
PR	30-MAR-2001; 2001US-0280186P.	
PR	01-MAY-2001; 2001US-0287735P.	
PR	05-JUN-2001; 2001US-0295848P.	
PR	25-JUN-2001; 2001US-0300465P.	
XX		
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.	
XX		
PI	Jackson DG, Feder J, Nelson T, Minter G, Ramanathan C, Lee L;	
PI	Stemmers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;	
PI	Krysatek S, Mcatee P, Suchard S, Banas D;	
DR	WPI, 2002-599721/64.	
XX	P-PsDB; ABR52361.	
PT	Novel polynucleotides encoding human phosphatase polypeptides useful in	
PT	the prevention or treatment of e.g. proliferative and cardiovascular	
PT	disorders.	
XX		
PS	Claim 1; Fig 13; 801p; English.	
XX		
CC	The invention relates to a novel isolated nucleic acid comprising a	

Query Match	100.0%	Score 906	DB 6	Length 5450
Best Local Similarity	100.0%	Pred. No. 7.8e-279		
Matches 906; Conservative	0	Mismatches 0	Indels 0	Cost 0

Qy	Db	Qy	Db
841	1318	901	1438
ATATCTCCAAATCTCAATTTCTGGGCAACTCTGACATGAGAAAGAAATTAGAAC	AAAGAGATGACATGCTTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAAGACCTCT	906	1443
1378	1378	906	1443
ATATCTCCAAATCTCAATTTCTGGGCAACTCTGACATGAGAAAGAAATTAGAAC	AAAGAGATGACATGCTTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAAGACCTCT	906	1443
Qy	Db	Qy	Db
901	1438	906	1443
CAGACT	CAGACT	906	1443
1438	1443	906	1443

ACCC60572
ID ACC60572 standard; cDNA; 5450 BP.
XX

13-JUN-2003 (First entry)
XX

INVENTION SEQ ID NO: 147.

antiparasitic; cardium; cytoskeletal; gene therapy; liver disease; proliferative disorder; renal failure; cardiovascular disorder; immunological disorder; arthritis; porriasis; congenital heart defect; congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly. Homo sapiens.

WO200257460-A1

25-JUL-2002.

20-DEC-2001; 2001WO-US050459.

20-DEC-2000; 2000US-0256868P.
30-MAR-2001; 2001US-0256868P.

01-MAY-2001; 2001US-0287735P.

25-JUN-2001; 2001US-0300465P

(BRIM) BRISTOL-MYERS SQUIBB CO

Jackson DG. Feder T. Welton T.

Siemens N, Bol D, Schieven G, Finger J, Toddernid CG, Baccollini C, Lee L; Krystek S, Mather J

WPT, 2000

P-PSDB; ABR52407.

Novel polynucleotides encoding human rheumatoid

disorders.

Example 1; Fig 19; 801pp; English.

polynucleotide relates to a novel isolated nucleic acid comprising a polynucleotide having a nucleotide sequence selected from 40 polynucleotides fully defined in the specification.

The invention has antiproliferative, hepatotropic, neurotrophic, antiatheritic, antiposrotic, cardiant, and cytostatic activity. The polynucleotide may have a use in gene therapy. A polynucleotide or ameliorating a medical condition, e.g. a proliferative disorder, or also useful for treating e.g. liver disease, renal failure, immunological disorders including arthritis and psoriasis, cardiovascular disorders such as congenital heart defects and congestive heart failure, and cancer. A method of the invention is useful for diagnosing a pathological condition or susceptibility to a condition in a subject. The present sequence is used in the exemplification of the invention.

1211 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;


```

Oy 61 AGTGAACGAAAAAGTGTCTGATATGATAGCCGGCCATTGTTGGAATACATACATCC 120
Db 85 AGTGAACGAAAAAGTGTCTGATATGATAGCCGGCCATTGTTGGAATACATACATCC 144
Oy 121 CACATTTTGGAAAGCCATTATATCAACTGCTCCAAAGTTTAAAGCGAAGTTGGCAACG 180
Db 145 CACATTTTGGAAAGCCATTATATCAACTGCTCCAAAGTTTAAAGCGAAGTTGGCAACG 204
Oy 181 GACAAAGTTTATATACAGAGCTCATCCAGCATTTAGGAAACATTAAGTTGACATTAT 240
Db 205 GACAAAGTTTATATACAGAGCTCATCCAGCATTTAGGAAACATTAAGTTGACATTAT 264
Oy 241 TGCAAGTGAAGAGTTGATTTAGATCAAGAGCTCCAAAGATGTTGCTCTCTCA 300
Db 265 TGCAAGTGAAGAGTTGATTTAGATCAAGAGCTCCAAAGATGTTGCTCTCTCTCA 324
Oy 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGAGTTCACTGTTCACTG 360
Db 325 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGAGTTCACTGTTCACTG 384
Oy 361 CTGTCAGGTGGTGTCTGAGATTTCTCTGCTGTTTCCCTGGCTCTGTAAGAAATCC 420
Db 385 CTGTCAGGTGGTGTCTGAGATTTCTCTGCTGTTTCCCTGGCTCTGTAAGAAATCC 444
Oy 421 ACTGAGTCCCTACCTGATTTCTGAGCCTTGCTTACCTGTTGCCAATGGGCAACC 480
Db 445 ACTGAGTCCCTACCTGATTTCTGAGCCTTGCTTACCTGTTGCCAATGGGCAACC 504
Oy 481 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGCAGATGTCCTCAACAAGAGCTGATA 540
Db 505 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGCAGATGTCCTCAACAAGAGCTGATA 564
Oy 541 CAGCAGATGGGATTTGTTATGTTAAATGCGACGTTTACTGTTCCAAAGCTGACTT 600
Db 565 CAGCAGATGGGATTTGTTATGTTAAATGCGACGTTTACTGTTCCAAAGCTGACTT 624
Oy 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 625 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684
Oy 661 CCGTGTGTCGACAAATCAGTATGATTTCAATTGAGAAAGAAAGCTCCCAATGATGTT 720
Db 685 CCGTGTGTCGACAAATCAGTATGATTTCAATTGAGAAAGAAAGCTCCCAATGATGTT 744
Oy 721 CTAGTGCATGTTTATGCTGGGATCTCCGCTCCGACACATGCTATGCTTACATCATG 780
Db 745 CTAGTGCATGTTTATGCTGGGATCTCCGCTCCGACACATGCTATGCTTACATCATG 804
Oy 781 AAGAGATGACATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 805 AAGAGATGACATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 864
Oy 841 ATATCTCCAAATCTCAATTTCTGGGCCAACTCTGAGCTATGAGAAAGATTTAGAAC 900
Db 865 ATATCTCCAAATCTCAATTTCTGGGCCAACTCTGAGCTATGAGAAAGATTTAGAAC 924
Oy 901 CAGACT 906
Db 925 CAGACT 930

```

RESULT 4

AAD09492 standard; DNA; 2732 BP.

AAD09492;

10-SEP-2001 (first entry)

Human SGP002 phosphatase polypeptide encoding DNA
 Human SGP002 phosphatase polypeptide; phosphatase-related disease;
 immune-related disorder; ocular disease; organ transplant rejection;

```

KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognition disorder; psychiatric disorder; cytostatic;
KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; vulnery; tranquiliser; antiepileptic;
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; mltaine; chromosome 12p11.1-p12.1; ds.
XX Homo sapiens.
XX OS
XX FH
XX Key Location/Qualifiers
XX CDS 538..2535
XX FT /*tag=a
XX FT /product="Human SGP002 phosphatase polypeptide"
XX MO200146394-A2.
XX PD 28-JUN-2001.
XX PF 21-DEC-2000; 2000MO-US034736.
XX PR 21-DEC-1999; 99US-0173255P.
XX PR 28-DEC-1999; 99US-0175766P.
XX PR 25-JAN-2000; 2000US-0178078P.
XX PR 31-JUN-2000; 2000US-0179301P.
XX (SUGEN-) SUGEN INC.
XX PA
XX PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;
XX Hill RJ, Flanagan P;
XX WPI; 2001-418058/44.
XX DR P-PSDB; AAE04834.
XX PT Novel phosphatase polypeptide useful for treating cancers, immune-related
XX diseases and disorders, cardiovascular disease, brain or neuronal-
XX associated diseases and metabolic disorders.
XX Claim 29; Fig 1; 186pp; English.
XX PS
XX CC The present invention relates to phosphatase polypeptides, nucleotide
XX sequences encoding them, as well as various products and methods useful
XX for the diagnosis and treatment of various phosphatase-related diseases
XX and conditions. Substance that modulates the activity of phosphatase
XX polypeptide is used to treat immune-related diseases and disorders,
XX cardiovascular disease, brain or neuronal-associated diseases and
XX metabolic disorders, including cancers of tissues, cancers of
XX haematopoietic origin, diseases of central and peripheral nervous system,
XX Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
XX lateral sclerosis, viral infections, infections caused by prions,
XX bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual
XX dysfunction, mood disorders, attention disorders, cognition disorders,
XX hypofunction, hypertension, psychotic disorders, neurological disorders,
XX dyskinesias and organ transplant rejection. The present sequence is a DNA
XX encoding human SGP002 phosphatase polypeptide. This sequence is
XX classified as dual specificity phosphatase (DSP) and MAP kinase
XX phosphatase (MKP). SGP002 gene maps to chromosomal position 12p11.1-p12.1
XX SQ

```

Query Match 99.6%; Score 902.8; DB 4; Length 2732;

Best Local Similarity 99.8%; Pred. No. 5.6e-278;

Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 1 ATGGCCATGAGATGATGGAACCAATTGTTACTGAGAGCTGTGCTGCTGGA 60
Db 538 ATGGCCATGAGATGATGGAACCAATTGTTACTGAGAGCTGTGCTGGA 597
Oy 61 AGTGAACGAAAAAGTGTCTGATATGATAGCCGGCCATTGTTGGAATACATACATCC 120

```

RESULT 3
AAF30479
ID AAF30479 standard; cDNA; 2118 BP.

	Query Match	Best Match	Local Similarity	Score	902.8;	DB 4;	Length	2118;
	Matches	904;	Conservative	0;	Mismatches	2;	Indels	0;
Qy	1	ATGGCCCATGAGTNGATTGAACTCAATTTGTTACTGAGAGGTGTGCTGTGGA	60					
Db	25	ATGGCCCATGAGTNGATTGAACTCAATTTGTTACTGAGAGGTGTGCTGTGGA	84					

Db 367 TGCAGTCAGAGGTTGATGTTACGATCAAGCTCCAGAGATGTGCTCTCTCTCA 426
Qy 301 GACTGTTTCTCACTGTAATCTTGGGTAACGTGAGAAAGCTTCAACTGTTCACCTG 360
Db 427 GACTGTTTCTCACTGTAATCTTGGGTAACGTGAGAAAGCTTCAACTGTTCACCTG 486
Qy 361 CTTCAGAGTGGGTTGCTGAGATCTCTGTTGTTTCCCTGGCCCTGTGAGAGAAATCC 420
Db 487 CTTCAGAGTGGGTTGCTGAGATCTCTGTTGTTTCCCTGGCCCTGTGAGAGAAATCC 546
Qy 421 ACTCTAGTCCCTACCTGATTTCTCAGCCCTGCTTACTCTGTGCAACATTTGGCCAA 480
Db 547 ACTCTAGTCCCTACCTGATTTCTCAGCCCTGCTTACTCTGTGCAACATTTGGCCAA 606
Qy 481 CGAATTTCTCCCAATCTTATCTTGGCTGCTGCAAGAGATGCTTCAACAGAGCTGATA 540
Db 607 CGAATTTCTCCCAATCTTATCTTGGCTGCTGCAAGAGATGCTTCAACAGAGCTGATA 666
Qy 541 CAGAGAAATGGGATTTGTTATGTTGTTAAATGCCAGCTTACTCTGCAAAAGCTGACTT 600
Db 667 CAGAGAAATGGGATTTGTTATGTTGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTT 726
Qy 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 727 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
Qy 661 CCGTGTGTGCAAAATCATGATGATTTCTTATGAGAAAGCAAAAGCTTCAATGATGTT 720
Db 787 CCGTGTGTGCAAAATCATGATGATTTCTTATGAGAAAGCAAAAGCTTCAATGATGTT 846
Qy 721 CTATGCACTGCTTATGCTGGGATCTCCCGCTCCCGCAATGCTATGCTTACATCATG 780
Db 847 CTATGCACTGCTTATGCTGGGATCTCCCGCTCCCGCAATGCTATGCTTACATCATG 906
Qy 781 AAGAGATGAGCATGCTTATGATGAGATGAGATTTGAGAAAGCAAAAGCTTACT 840
Db 907 AAGAGATGAGCATGCTTATGATGAGATGAGATTTGAGAAAGCAAAAGCTTACT 966
Qy 841 ATATCTCCAAATCTCAATTTCTGCGCAACTCTCTGAGCTATGAGAGATTAAGAAC 900
Db 967 ATATCTCCAAATCTCAATTTCTGCGCAACTCTCTGAGCTATGAGAGATTAAGAAC 1026
Qy 901 CAGACT 906
Db 1027 CAGACT 1032

RESULT 6
ABK47596
ID ABK47596 standard; cDNA, 3496 BP.
XX
AC ABK47596;
XX
DT 02-JUL-2002 (first entry)
XX
DE cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.
XX
KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contract inhibition;
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 562..2559
FT /tag=a
FT /product="Human dual-specificity phosphatase-3 (DSP-16)
FT protein"

PN WO200226997-A2.
XX 04-APR-2002.
XX
XX 25-SEP-2001; 2001MO-US030124.
XX
XX 26-SEP-2000; 2000US-0235487P.
XX
XX (CEPT-) CEPTIR INC.
XX
XX Luche RM, Wei B;
XX
XX WPI: 2002-315802/35.
XX P-PsDB; AAU79156.
XX
XX New DSP-16 polypeptide, useful for identifying modulators of its
XX activity, which can be used in the treatment of disorders such as
XX Duchenne muscular dystrophy, or cancer.
XX
XX Claim 7; Fig 1; 87pd; English.
XX
XX The present invention relates to a new polypeptide, DSP-16, having a 665
XX amino acid sequence, given in the specification, or a variant having at
XX least 50 % identical residues, which retains the ability to
XX dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
XX invention can be used for identifying agents which modulate DSP-16
XX activity for modulation of a proliferative response in a cell, survival
XX of a cell, or differentiation of a cell. The cell displays contact
XX inhibition of cell growth or anchorage independent growth and may display
XX altered intercellular adhesion. The agent may modulate apoptosis, or the
XX cell cycle. The identified modulators can be used to treat Duchenne
XX muscular dystrophy, cancer, graft-versus-host disease, autoimmune
XX diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
XX cell proliferation, and cell cycle abnormalities. The present nucleic
XX acid sequence is that of the human dual-specificity phosphatase-3 (DSP-
XX 16) gene located on chromosome 12p. This sequence encodes the human DSP-
XX 16 protein of the invention

SO Sequence 3496 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;
Query Match 99.6%; Score 902.8; DB 6; Length 3496;
Best Local Similarity 99.8%; Pred. No. 6.5e-278;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCCCATGAGATGATGAGATGAGATGATGAGATGAGATGAGATGAGATGAGATGAG 60
Db 562 ATGGCCCATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 621
Qy 61 AGTGAACGAGAAAGAGTCTGCTAATGATGATGATGATGATGATGATGATGATGATGAT 120
Db 622 AGTGAACGAGAAAGAGTCTGCTAATGATGATGATGATGATGATGATGATGATGATGAT 681
Qy 121 CACATTTGAGAGCATTAATATCACTGCTCCAGCTTATGAGAGCAAGTTGCCAAG 180
Db 682 CACATTTGAGAGCATTAATATCACTGCTCCAGCTTATGAGAGCAAGTTGCCAAG 741
Qy 181 GACAAAGTGTAAATTAACAAGAGCTATGAGAGAGCTATGAGAGAGCTATGAGAGAGCT 240
Db 742 GACAAAGTGTAAATTAACAAGAGCTATGAGAGAGCTATGAGAGAGCTATGAGAGAGCT 801
Qy 241 TGCAGTCAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 802 TGCAGTCAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 861
Qy 301 GACTGTTTCTCACTGTAATCTTGGGTAACGTGAGAAAGCTTCAACTGTTCACCTG 360
Db 862 GACTGTTTCTCACTGTAATCTTGGGTAACGTGAGAAAGCTTCAACTGTTCACCTG 921
Qy 361 CTTCAGAGTGGGTTGCTGAGATCTCTGTTGTTTCCCTGGCCCTGTGAGAGAAATCC 420
Db 922 CTTCAGAGTGGGTTGCTGAGATCTCTGTTGTTTCCCTGGCCCTGTGAGAGAAATCC 981
Qy 421 ACTCTAGTCCCTACCTGATTTCTCAGCCCTGCTTACTCTGTGCAACATTTGGCCAA 480

Db	598	ACGTGAACGGAAAAAGCTGCTCTAAATTGATAACCGGACCATTTGTGGATAACATAACATCC	657
OY	121	CACATTTTGGAGCCATTAAATATCAACTGCTCCAGACTTATGAAGCGAAGGTTGCAACG	180
Db	658	CACATTTTGGAGCCATTAAATATCAACTGCTCCAGACTTATGAAGCGAAGGTTGCAACG	717
OY	181	GACAAAGTGTAAATATACAGAGCTCATCCAGAAATTCAGGAAACATAAGCTTGACATGGAT	240
Db	718	GACAAAGTGTAAATATACAGAGCTCATCCAGAAATTCAGGAAACATAAGCTTGACATGGAT	777
OY	241	TGCGAGTGAAGGTTGTATGTTTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA	300
Db	778	TGCGAGTGAAGGTTGTATGTTTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA	837
OY	301	GACTGTTTTCACCTGATCTCTGGGTAAACTGGAGAAAGGCTTCAACTCTGTACCTG	360
Db	838	GACTGTTTTCACCTGATCTCTGGGTAAACTGGAGAAAGGCTTCAACTCTGTACCTG	897
OY	361	CTTGCAAGTGGGTTTGCTGAGTCTCTCGTTGTTTCCCTGGCCTCTGTAGAGAAATCC	420
Db	898	CTTGCAAGTGGGTTTGCTGAGTCTCTCGTTGTTTCCCTGGCCTCTGTAGAGAAATCC	957
OY	421	ACTCTAGTCCCTAACCTGACATTTCTCAAGCCTTGCTTACCTGTGGCAACATTTGGGCCAAC	480
Db	958	ACTCTAGTCCCTAACCTGACATTTCTCAAGCCTTGCTTACCTGTGGCAACATTTGGGCCAAC	1017
OY	481	CGAATCTTCCCAATCTTATCTTGGCTGCACGCGAGATGTCTTCAACAAAGAGCTGATA	540
Db	1018	CGAATCTTCCCAATCTTATCTTGGCTGCACGCGAGATGTCTTCAACAAAGAGAGCTGATG	1077
OY	541	CAGCAGAAATGGAGATTGGTATGTGTAAATGCGAGCTATACCTGATCCCAAGCTGACCTT	600
Db	1078	CAGCAGAAATGGAGATTGGTATGTGTAAATGCGAGCTATACCTGATCCCAAGCTGACCTT	1137
OY	601	ATCCCCGAGCTCATTTCTCGCGTGGCCCTGTGAATGACAGCTTTTGTGAGAAAAATTTG	660
Db	1138	ATCCCCGAGCTCATTTCTCGCGTGGCCCTGTGAATGACAGCTTTTGTGAGAAAAATTTG	1197
OY	661	CCGTGTGGACAAATCAGATAGATTTCATTGAGAAAGCAAAAGCCTCCAAATGAGATGGTT	720
Db	1198	CCGTGTGGACAAATCAGATAGATTTCATTGAGAAAGCAAAAGCCTCCAAATGAGATGGTT	1257
OY	721	CTAGTGCACCTGTTTGTGCTGGGATCTCCGCGTCCGCGCAGCATGCTATGCGCTTACATCATG	780
Db	1258	CTAGTGCACCTGTTTGTGCTGGGATCTCCGCGTCCGCGCAGCATGCTATGCGCTTACATCATG	1317
OY	781	AAGAGGATGCAACATGCTTTTATGATGAAGACTTACAGTTTGTGAAAAGAAAAAGACCTACT	840
Db	1318	AAGAGGATGCAACATGCTTTTATGATGAAGACTTACAGTTTGTGAAAAGAAAAAGACCTACT	1377
OY	841	ATATCTCCAAACTCAATTTTCTGGGCAACTCTCGACATATGAAAGAAAGATTAAAGAAC	900
Db	1378	ATATCTCCAAACTCAATTTTCTGGGCAACTCTCGACATATGAAAGAAAGATTAAAGAAC	1437
OY	901	CAGACT 906	
Db	1438	CAGACT 1443	

RESULT 5
AAS15768
ID AAS15768 standard; cDNA; 3059 BP
XX

DT 18-JUN-2002 (first entry)
XX

DE
XX

cDNA sequence of human (dual specificity phosphatase) DUSP-10.

KW Human; dual specificity phosphatase, DUSP-10; cancer; epilepsy; stroke;
KW neuronal degeneration syndrome; Alzheimer's disease; depression;
KW schizophrenia; asthma; immune disorder; inflammatory process; arthritis
KW osteoporosis; diabetes; gene therapy; chromosome 12; 58.

XX	OS	Homo sapiens.
XX	XX	
XX	Key	Location/Qualifiers
XX	CDS	127..2124
XX	FT	/*tag= a
XX	FT	/product= "DUSP-10 protein"
XX	XX	
XX	XX	MO200177340-A1.
XX	PD	18-OCT-2001.
XX	XX	
XX	PF	06-APR-2001; 2001WO-EP003966.
XX	XX	
XX	PR	10-APR-2000; 2000EP-00107143.
XX	PA	
XX	XX	(MERE) MERCK PATENT GMBH.
XX	PI	
XX	DR	Duecker K;
XX	DR	WPI; 2002-010917/01.
XX	XX	P-PSDB; AAU09946.
XX	XX	
XX	XX	Novel dual specificity phosphatase polypeptides useful for treating
XX	XX	cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
XX	XX	disease, depression, schizophrenia, asthma and immune disorders.
XX	XX	Claim 5; Page 34-37; 43pp; English.

the present invention relates to a new isolated dual specificity phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid sequence that is fully defined in the specification. The invention also provides a sequence encoded by a 3059 nucleotide sequence fully defined in the specification, and a sequence having at least 95 % identity to the polypeptide, or fragments or variants of DUSP-10. The invention is useful for treating cancer e.g. leukaemia, colon carcinoma, lung cancer, prostate cancer, metastasis of tumour cells, neuro-angiodysgenesis, epilepsy, stroke, neuronal degeneration of synovial cells, Alzheimer's disease, depression, schizophrenia, cardiac myopathies, asthma, immune disorders, inflammatory processes e.g. arthritis, bowel disease, type I diabetes, osteoporosis, diabetes and diabetes associated diseases. The molecules of the invention are also useful as vaccines for inducing immunological response in a mammal, in disease diagnosis and in assays for screening agonistic or antagonistic compounds. Other uses of the invention include identifying membrane bound or soluble receptors, as a diagnostic reagent, in chromosome localisation studies, and as a valuable tool in tissue expression studies. The present sequence represents cDNA of the human dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome 12

Query Match	99.6%	Score 902.8	DB 6	Length 3059
Best Local Similarity	99.8%	Pred. No. 6e-278		
Matches 904; Conservative	0	Mismatches 2	Indels 0	Gaps 0

QY	ATGGCCCATGAGATGATTGGAACTCAAAATTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA	60
Dp	127 ATGGCCCATGAGATGATTGGAACTCAAAATTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA	186
QY	61 AGTGAACGGAAAAAGTCTGCTAATTGATGACGGGCACTTGTGGAAATACAAATACATCC	120
Dp	187 AGTGAACGGAAAAAGTCTGCTAATTGATGACGGGCACTTGTGGAAATACAAATACATCC	246
QY	121 CACATTTTGGAAAGCCATTAAATATATCAATGCTCCAAAGTTATGAAGCGAAGTTGGCAACAG	180
Dp	247 CACATTTTGGAAAGCCATTAAATATCAATGCTCCAAAGTTATGAAGCGAAGTTGGCAACAG	306
QY	181 GACAAAGTGTAAATTACAGAGCTATCCAGCAATTCAGCGAAACATTAAGTTGACATTTGAT	240
Dp	307 GACAAAGTGTAAATTACAGAGCTATCCAGCAATTCAGCGAAACATTAAGTTGACATTTGAT	366
QY	241 TGCAGTCAGAAAGTTGATGTTACGATCAAAGCTCCCAAGATGTGGCTCTCTCTCTTCA	300

Qy	601	ATCCCGAGTCTCAATTTCTCGGTGCTGGAAATGACAGTTTGTGAGAAAATTGG	660
Db	1189	ATCCCGAGTCTCAATTTCTCGGTGCTGGAAATGACAGTTTGTGAGAAAATTGG	1248
Qy	661	CCGTGTTGGAACAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCMATGATGTGT	720
Db	1249	CCGTGTTGGAACAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCMATGATGTGT	1308
Qy	721	CTAGTGCATGTTTAGTCTGGGATCTCCGCTCCGCCACATGCGTATGCGCTACATCATG	780
Db	1309	CTAGTGCATGTTTAGTCTGGGATCTCCGCTCCGCCACATGCGTATGCGCTACATCATG	1368
Qy	781	AAGAGATGAGACATGCTTTTAGATGTAACCTTACAGATTGTGGAAGAAAAAGACTTACT	840
Db	1369	AAGAGATGAGACATGCTTTTAGATGTAACCTTACAGATTGTGGAAGAAAAAGACTTACT	1428
Qy	841	ATATCTCAAACTTCAATTTTCTGGGCAACCTCCGGACTATGAGAGAAGATTAAAGC	900
Db	1429	ATATCTCAAACTTCAATTTTCTGGGCAACCTCCGGACTATGAGAGAAGATTAAAGC	1488
Qy	901	CAGACT 906	
Db	1489	CAGACT 1494	

RESULT 8
 ABK49402
 ID ABK49402 standard; CDNA; 3544 BP.
 XX
 AC ABK49402;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE cDNA encoding human dual specificity phosphatase 21117 protein.
 XX
 KW Human; dual specificity phosphatase 21117; erythroid-related disorder;
 KW haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
 KW erythrocytosis; liver-related disorder; cancer; gene; ss.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 CDS 589..2586
 FT /*tag= a
 FT /product= "Human dual specificity phosphatase 21117"
 FT /note= "Specifically claimed in claim 2"
 XX
 PN US2002034807-A1.
 XX
 PD 21-MAR-2002.
 XX
 PF 23-MAR-2001; 2001US-00816494.
 XX
 PR 24-MAR-2000; 2000US-0191858P.
 XX
 PA (MEYE/) MEYERS R A.
 XX
 PI Meyers RA;
 XX
 DR MPI; 2002-351088/36
 DR P-PSDB; AAU79929
 XX
 PT New nucleic acids designated 36692 and 21117, encoding dual specificity
 PT phosphatases for treating cell proliferation and differentiation
 PT disorders including hematopoietic and erythroid-related disorders and
 PT cancers.
 XX
 PS Claim 2; Fig 1; 76pp; English.
 XX
 CC The present invention relates to new nucleic acids designated 36692 and
 CC 21117 encoding dual specificity phosphatase family members. The nucleic
 CC acid, polypeptide encoded by it, and antibody specific for the

CC polypeptide may be used to diagnose and treat haematopoietic-related
CC disorders such as leukemias and autoimmune diseases, erythroid-related
CC disorders such as anaemias and erythrocytosis, liver-related disorders,
CC and cancers, particularly of the breast, colon, adipose, prostate and
CC lung. The present nucleic acid sequence encodes the human dual
CC specificity phosphatase 2117 protein of the invention, as described
CC above
CC
CC
CC
CC Sequence 3544BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
XX

Query Match	99.6%;	Score	502.8;	DB	6;	Length	3544;
Best Local Similarity	99.6%;	Pred.	No. 6.5e-278;				
Matches	904;	Conservative	0;	Mismatches	2;	Indels	0;
QY	1	ATGGCCCATGATGATGTGGAACTCAAAATTGTTACTGAGAGGTGTGGCTCTGCTGAA	60				
DB	589	ATGGCCCATGATGATGTGGAACTCAAAATTGTTACTGAGAGGTGTGGCTCTGCTGAA	648				
QY	61	AGTGGAAACGGAAAAAGTGTCTCTAATTGATATAGCCGGGCAATTTGTGGAATCAATCAATCC	120				
DB	649	AGTGGAAACGGAAAAAGTGTCTCTAATTGATATAGCCGGGCAATTTGTGGAATCAATCAATCC	708				
QY	121	CACATTTTGGAGCCATTAATATCAACTGTCCCAAGCTTATGAAAGCGAAGTTGCAACAG	180				
DB	709	CACATTTTGGAGCCATTAATATCAACTGTCCCAAGCTTATGAAAGCGAAGTTGCAACAG	768				
QY	181	GACAAAGTGTAAATTAACAGAGCTCATCAGACATTCAGCGAAACATAAGTTGACATTGAT	240				
DB	769	GACAAAGTGTAAATTAACAGAGCTCATCAGACATTCAGCGAAACATAAGTTGACATTGAT	828				
QY	241	TGCAGTCAGAAAGTTGATGATTACGATCAAAGCTCCCAAGATGTGTGCTCTCTCTTCA	300				
DB	829	TGCAGTCAGAAAGTTGATGATTACGATCAAAGCTCCCAAGATGTGTGCTCTCTCTCTTCA	888				
QY	301	GACGTGTTTTCACGTACTCTTGCGGTTAACTGAGAAAGACTTCAACTCTGTTCACCTG	360				
DB	889	GACGTGTTTTCACGTACTCTTGCGGTTAACTGAGAAAGACTTCAACTCTGTTCACCTG	948				
QY	361	CTTGACAGTGGGTTTGGCTGAATTCTCTGTTGTTCCCTGGCTCTGTGAAAGAAAAATCC	420				
DB	949	CTTGACAGTGGGTTTGGCTGAATTCTCTGTTGTTCCCTGGCTCTGTGAAAGAAAAATCC	1008				
QY	421	ACTTAAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCCTGTGGCAACATTGGGCCCAAC	480				
DB	1009	ACTTAAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCCTGTGGCAACATTGGGCCCAAC	1068				
QY	481	CGAATTCCTCCCAATCTTTAATCTTGGGTGCGCAGGAGATGTGCTCAACAAGAGCTGATA	540				
DB	1069	CGAATTCCTCCCAATCTTTAATCTTGGGTGCGCAGGAGATGTGCTCAACAAGAGCTGATA	1128				
QY	541	CAGCAGATGGGATGGTATATGTGTTAAATGCGCAGCTAATACCTGTCCAAAGCTGACTTT	600				
DB	1129	CAGCAGATGGGATGGTATATGTGTTAAATGCGCAGCTAATACCTGTCCAAAGCTGACTTT	1188				
QY	601	ATCCCCGAGTCTCAATTCCTGCTGCTGCTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG	660				
DB	1189	ATCCCCGAGTCTCAATTCCTGCTGCTGCTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG	1248				
QY	661	CCGTGCTTTGGAACAATCAGTAGATTTTCAATGAGAAAGCAAAAGGCTCCAAATGAGATGTT	720				
DB	1249	CCGTGCTTTGGAACAATCAGTAGATTTTCAATGAGAAAGCAAAAGGCTCCAAATGAGATGTT	1308				
QY	721	CTAATGCACTGTTTAGCTGGGATCTCCGCGCTCCGCAACATCGCTATGCGCTCAATCATG	780				
DB	1309	CTAATGCACTGTTTAGCTGGGATCTCCGCGCTCCGCAACATCGCTATGCGCTCAATCATG	1368				
QY	781	AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTTACT	840				
DB	1369	AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTTACT	1428				
QY	841	ATATCTCCAACTTCATTTTCTTGCGGCAACTCTGTGACATATGGAAGAAATTAAGAAC	900				
DB	1429	ATATCTCCAACTTCATTTTCTTGCGGCAACTCTGTGACATATGGAAGAAATTAAGAAC	1488				

```
Db 982 ACTTAGTCCCTACCTGATTTCTCAGCTTGTCTTACTCTTGTCCAACTTGGCCAACTC 1041
QY 481 CGAATTTCTCCCAATCTTTAATCTTGGCTGCGCAGCGAGATGCTCTCAACAAGAGCTGATA 540
Db 1042 CGAATTTCTCCCAATCTTTAATCTTGGCTGCGCAGCGAGATGCTCTCAACAAGAGCTGATA 540
QY 541 CAGCAGATGGAGTTGGTTATGTATGTAAATGACGATATACCTGTCCAAAGCTGACTTT 600
Db 1102 CAGCAGATGGAGTTGGTTATGTATGTAAATGACGATATACCTGTCCAAAGCTGACTTT 600
QY 601 ATCCCGAGTCTATTTCTTGGCTGCTGCTGGAATGACAGCTTTGTGAGAAAATTTTG 660
Db 1162 ATCCCGAGTCTATTTCTTGGCTGCTGCTGGAATGACAGCTTTGTGAGAAAATTTTG 660
QY 661 CCGTGTGGACAAATCAGTATGTTCAATGAGAAAGCAAGCTCCAAATGATGTTGT 720
Db 1222 CCGTGTGGACAAATCAGTATGTTCAATGAGAAAGCAAGCTCCAAATGATGTTGT 720
QY 721 CTAGTCACTGTTTGTGGCTGCGCTCCGCTCCGCAACATGCTATGCTCAATGATG 780
Db 1282 CTAGTCACTGTTTGTGGCTGCGCTCCGCTCCGCAACATGCTATGCTCAATGATG 780
QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGTGAAAAGAAAAGAGCTTACT 840
Db 1342 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGTGAAAAGAAAAGAGCTTACT 840
QY 841 ATATCTCAAACTTCAATTTTCTGGCCAACTCCTGACATGAGAAAGATTAAGAAC 900
Db 1402 ATATCTCAAACTTCAATTTTCTGGCCAACTCCTGACATGAGAAAGATTAAGAAC 900
QY 901 CAGACT 906
Db 1462 CAGACT 1467
```

RESULT 7
AAS14639
ID AAS14639 standard; cDNA; 3544 BP.

AC AAS14639;
DT 18-DEC-2001 (first entry)
XX

DE Human cDNA encoding dual specificity phosphatase 21117.

XX Human; ss; dual specificity phosphatase 21117; hepatotropic; cytosolic;
KW rheumatoid arthritis; autoimmune disorder; diabetes mellitus;
KW liver disorder; erythroid associated disorder; haemolytic anaemia;
KW cellular proliferative disorder; leukaemia;
KW acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer.

OS Homo sapiens.

XX Key Location/Qualifiers
FT CDS 589..2586
FT /*cag= a
FT /product= "Dual specificity phosphatase 21117"

PN MO200173059-A2.

PD 04-OCT-2001.

PF 23-MAR-2001; 2001MO-US009477.

PR 24-MAR-2000; 2000US-0191858P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Meyers RA;

DR WPI; 2001-611635/70.

```
DR P-PSDB; AAU09016.
XX New human dual specificity polypeptides and nucleic acids for diagnosis
PT of disease and treatment of e.g. liver disorders.
XX Claim 1; Fig 1; 143pb; English.

The invention relates to two novel human dual specificity phosphatases
designated 21117 and 38692, the nucleic acids encoding them (including
fragments, allelic variants, their complements or nucleic acids that
hybridise to them) and antibodies raised against the proteins. The
antibody is useful for detecting the presence of the polypeptide, and the
nucleic acid fragments are useful for detecting the presence of the
nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
antisense sequences) are useful for modulating the activity of expression
of the polypeptide/nucleic acid, useful for the treatment of e.g.
rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
disorders (e.g. haemolytic anaemia) cellular proliferative or
differentiative disorders, leukaemias (e.g. acute myeloid leukaemia),
carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
38692 are also useful for modulating the proliferation, survival, and
migration or differentiation of a 38692 or 21117-expressing cell. The
polypeptide and nucleic acids are useful for identifying modulating
agents. The present sequence encodes the dual specificity phosphatase
21117.

Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;

Query Match 99.6%; Score 902.8; DB 5; Length 3544;
Best Local Similarity 99.8%; Pred. No. 6; 5e-278;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCCCATGAGATGATGTAACCAATGTTACTGAGAGTGGCTGCTGCTGGA 60
Db 589 ATGCCCATGAGATGATGTAACCAATGTTACTGAGAGTGGCTGCTGCTGGA 648
QY 61 AGTGAACGAAAAAGTGTCTAATGATGAGCGGSCATTTGTGAATATATATATCC 120
Db 649 AGTGAACGAAAAAGTGTCTAATGATGAGCGGSCATTTGTGAATATATATATCC 708
QY 121 CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAGCGAGTGGCAACAG 180
Db 709 CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAGCGAGTGGCAACAG 768
QY 181 GACAAAGTGTAAATTCAGAGCTCATCCAGCATTCAGAGGAAACATAAGTTGACATTGAT 240
Db 769 GACAAAGTGTAAATTCAGAGCTCATCCAGCATTCAGAGGAAACATAAGTTGACATTGAT 828
QY 241 TGCAGTCAGAAAGTTGTATGATGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 300
Db 829 TGCAGTCAGAAAGTTGTATGATGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 888
QY 301 GACTGTTTCTCACTGATCTTCTGGTAAACTGGAAGAGGCTCAACTGTTCACTTG 360
Db 889 GACTGTTTCTCACTGATCTTCTGGTAAACTGGAAGAGGCTCAACTGTTCACTTG 948
QY 361 CTTGACGATGGGTTTGTGAGATTCTCTGTTGTTTCTCTGGCCCTCTGTGAAGAAAATCC 420
Db 949 CTTGACGATGGGTTTGTGAGATTCTCTGTTGTTTCTCTGGCCCTCTGTGAAGAAAATCC 1008
QY 421 ACTCTAGTCCCTACCTCATTTCTCAGCCCTGTTACCTGTTGCCAACTTGGCCAACTC 480
Db 1009 ACTCTAGTCCCTACCTCATTTCTCAGCCCTGTTACCTGTTGCCAACTTGGCCAACTC 1068
QY 481 CGAATTTCTCCCAATCTTTAATCTTGGCTGCGCAGCGAGATGCTCTCAACAAGAGCTGATA 540
Db 1069 CGAATTTCTCCCAATCTTTAATCTTGGCTGCGCAGCGAGATGCTCTCAACAAGAGCTGATA 540
QY 541 CAGCAGATGGAGTTGGTTATGTATGTAAATGACGATATACCTGTCCAAAGCTGACTTT 600
Db 1129 CAGCAGATGGAGTTGGTTATGTATGTAAATGACGATATACCTGTCCAAAGCTGACTTT 1188
```

```

RESULT 10
ABN83966
ID ABN83966 standard; DNA; 4790 BP.
XX
AC ABN83966;
XX
DT 06-SEP-2002 (first entry)
XX
DE Human gene sequence #13.
XX
KW Human; brain; tonsil; hippocampus; foetal brain; diagnosis; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 184..2181
FT FT /*tag= a
XX
XX WO200252005-A1.
XX
XX 04-JUL-2002.
XX
XX 20-DEC-2001; 2001WO-JP011217.
XX
XX 22-DEC-2000; 2000JP-00389742.
XX
XX (KAZU-) KAZUSA DNA RES INST FOUND.
XX
XX (CELE-) CELESTAR LEXICO-SCI LTD.
XX
XX Ohara O, Negase T, Nakajima D;
XX
XX WPI; 2002-500762/53.
XX
XX P-PSDB; ABB97946.
XX
XX Genes and their expression products cloned from human cDNA libraries for
XX treatment and diagnosis of diseases associated with their expression.
XX
XX Claim 1(a); Page 111-117; 238pp; Japanese.
XX
XX The invention relates to DNA encoding polypeptides directly cloned from
XX cDNA libraries originating in adult whole brain, human tonsil, human
XX adult hippocampus and human foetal whole brain. Polypeptides and
XX polynucleotides of the invention may be used in the investigation of
XX differential expression of the DNA sequences in normal subjects and
XX disease patients. They may also be used in the production of antibodies,
XX oligonucleotide probes and DNA chips for diagnosis and identification of
XX drugs for treatment of diseases with which the DNA sequences are
XX associated. The sequences given in records ABN83954-ABN83984 represent
XX human gene sequences of the invention
XX
XX SQ Sequence 4790 BP; 1293 A; 1064 C; 1152 G; 1281 T; 0 U; 0 Other;
XX
XX Query Match 99.6%; Score 902.8; DB 6; Length 4790;
XX Best Local Similarity 99.8%; Pred. No. 7.7e-278;
XX Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 ATGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGGTTGGTGCTCTCTCGAA 60
DB 184 ATGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGGTTGGTGCTCTCTCGAA 243
QY 61 AGTGGACCGGAAAAGTCTCTCTAAATGATAGCGGCCCATTTGGGAATACAATACATCC 120
DB 244 AGTGGACCGGAAAAGTCTCTCTAAATGATAGCGGCCCATTTGGGAATACAATACATCC 303
QY 121 CACATTTTGGAGCCATTAATATCACTGCTCCAAAGCTTATGAAGCGAGGTTGCAACAG 180
DB 304 CACATTTTGGAGCCATTAATATCACTGCTCCAAAGCTTATGAAGCGAGGTTGCAACAG 363
QY 181 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAACAATAAGGTTGACATTGAT 240
DB 364 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAACAATAAGGTTGACATTGAT 423
QY 241 TGCAGTCAGAAAGGTTGTAGTTACGATCAAAAGCTCCCAAGATGTTGCGCTCTCTCTTCA 300

```

RESULT 11

ABV20833

ID ABV20833 standard; cDNA; 5145 BP.

XX

AC ABV20833;

XX

DT 13-SEP-2002 (first entry)

XX

Human prostate expression marker cDNA 20824.

DE

XX

XX

KW

XX

OS

XX

XX

PN

XX

PD

XX

XX

PF

XX

PR

XX

PR

XX

PR

XX

PR

XX

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.

Homo sapiens.

WO200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US005171.

17-FEB-2000; 2000US-0183319P.

16-MAR-2000; 2000US-0189862P.

25-MAY-2000; 2000US-0207454P.

09-JUN-2000; 2000US-0211314P.

QY 901 CAGACT 906
DB 1489 CAGACT 1494
RESULT 9
ABK14474
ID ABK14474 standard; cDNA; 3766 BP.
XX AC ABK14474;
XX DT 08-MAY-2002 (first entry)
XX DE Human protein phosphatase 7 (pp7) cDNA sequence.
XX KW Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;
KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
KW melanoma; myeloma sarcoma; gene; ss.
XX OS Homo sapiens.
XX FH Key
XX FT Location/Qualifiers
XX FT 538..2535
XX FT CDS
XX FT /*tag= a
XX FT /product= "Protein_phosphatase_7 (pp7)"
XX PN WO200210363-A2.
XX PD 07-FEB-2002.
XX PF 26-JUL-2001; 2001WO-US023716.
XX PR 28-JUL-2000; 2000US-0221679P.
XX PR 03-AUG-2000; 2000US-0223272P.
XX PR 10-AUG-2000; 2000US-0224309P.
XX PR 18-AUG-2000; 2000US-0226728P.
XX PR 30-AUG-2000; 2000US-0229254P.
XX PR 08-SEP-2000; 2000US-0231366P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;
PI Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalia AJA, Lu DAM;
PI Tribouley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
PI Wallia NK, Kearney LJ;
XX DR WPI; 2002-188735/24.
XX DR P-PSDB; AAU75789.
XX PT New protein phosphatases, useful for diagnosing, treating or preventing
XX PT immune system disorders (e.g. Crohn's disease), neurological disorders
XX PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
XX PT cancers).
XX PS Claim 5; Page 114-115; 117pp; English.
XX CC The present invention relates to a new polypeptide, a naturally occurring
XX CC amino acid sequence at least 95% identical to it, a biologically active
XX CC fragment of it or an immunogenic fragment of it. The polypeptides,
XX CC polynucleotides, agonists and antagonists are useful for diagnosing,
XX CC treating or preventing disorders associated with aberrant expression of
XX CC protein phosphatases (PP), particularly immune system disorders e.g.
XX CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
XX CC asthma or Crohn's disease, neurological disorders e.g. epilepsy,
XX CC Huntington's disease, dementia or Parkinson's disease, developmental
XX CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
XX CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
XX CC or sarcoma. The present nucleic acid sequence encodes human protein
XX CC phosphatase 7 (pp7) which is one of several human protein phosphatases

CC (AAU75783-AAU75792) of the invention
XX
SQ Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;
Query Match 99.6%; Score 902.8; DB 6; Length 3766;
Best Local Similarity 99.8%; Pred. No. 6.7e-278;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGCCCATGAGATGATTGGAATCTCAAAATGTTTACTGAGAGGTTGGTGGCTCTGCTGGAA 60
DB |||||
QY 538 ATGCCCATGAGATGATTGGAATCTCAAAATGTTTACTGAGAGGTTGGTGGCTCTGCTGGAA 597
DB |||||
QY 61 AGTGGAAACGGAAGAGTGTCTTAATTGATAGCGGCCCAATTTGTGGAATACAAATACATCC 120
DB |||||
QY 598 AGTGGAAACGGAAGAGTGTCTTAATTGATAGCGGCCCAATTTGTGGAATACAAATACATCC 657
DB |||||
QY 121 CACATTTTGGAGCCCAATTAATCACTCTCAAGCTTATCAAGCAAGGTTGCAACAG 180
DB |||||
QY 658 CACATTTTGGAGCCCAATTAATCACTCTCAAGCTTATCAAGCAAGGTTGCAACAG 717
DB |||||
QY 181 GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGCGAAACATAAGGTTGCAATTCAT 240
DB |||||
QY 718 GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGCGAAACATAAGGTTGCAATTCAT 777
DB |||||
QY 241 TGCAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCTTCA 300
DB |||||
QY 778 TGCAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCTTCA 837
DB |||||
QY 301 GACTGTTTCTCAGCTGTACTTCTGGGTAAACTGGAGAGAGCTTCAACTCTGTTCACTG 360
DB |||||
QY 838 GACTGTTTCTCAGCTGTACTTCTGGGTAAACTGGAGAGAGCTTCAACTCTGTTCACTG 897
DB |||||
QY 361 CTTGCAGCTGGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGCTCTCTGTAAGGAAATCC 420
DB |||||
QY 898 CTTGCAGCTGGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGCTCTCTGTAAGGAAATCC 957
DB |||||
QY 421 ACTTAGTCCCTACTCCTGATTTCTCAGCTTTGCTTACTCTGTTGCCAATTTGGGCCAACC 480
DB |||||
QY 958 ACTTAGTCCCTACTCCTGATTTCTCAGCTTTGCTTACTCTGTTGCCAATTTGGGCCAACC 1017
DB |||||
QY 481 CGAATTTCTCCAAATCTTTTATCTTGGCTGCCAGAGATGTCCTCAACAAGGAGCTGATA 540
DB |||||
QY 1018 CGAATTTCTCCAAATCTTTATCTTGGCTGCCAGAGATGTCCTCAACAAGGAGCTGATG 1077
DB |||||
QY 541 CAGCAGATGGGATTTGGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 600
DB |||||
QY 1078 CAGCAGATGGGATTTGGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTT 1137
DB |||||
QY 601 ATCCCAGTCTCATTTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
DB |||||
QY 1138 ATCCCAGTCTCATTTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1197
DB |||||
QY 661 CCGTGGTTGGACAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCCTTCAATGGATGTT 720
DB |||||
QY 1198 CCGTGGTTGGACAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCCTTCAATGGATGTT 1257
DB |||||
QY 721 CTAGTGCATGTTTACGTGGGATCTCCCGCTCGGCCCATCGCTATCGCTACATCATG 780
DB |||||
QY 1258 CTAGTGCATGTTTACGTGGGATCTCCCGCTCGGCCCATCGCTATCGCTACATCATG 1317
DB |||||
QY 781 AAGAGCATGACATGCTTTTAGATGAGCTTACAGATTTGTGAAAGAAAAGACCTACT 840
DB |||||
QY 1318 AAGAGCATGACATGCTTTTAGATGAGCTTACAGATTTGTGAAAGAAAAGACCTACT 1377
DB |||||
QY 841 ATATCTCCAAATCTCAATTTTCTGGGCCAACTCTCGGACTATGAGAAAGCAATTAAGAAC 900
DB |||||
QY 1378 ATATCTCCAAATCTCAATTTTCTGGGCCAACTCTCGGACTATGAGAAAGCAATTAAGAAC 1437
DB |||||
QY 901 CAGACT 906
DB 1438 CAGACT 1443

PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JR;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3419; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Query Match 99.64; Score 902.8; DB 5; Length 5145;
Best Local Similarity 99.84; Pred. No. 8e-278; Indels 0; Gaps 0;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGTGAATTGGAACCAATTGTACTGAGAGTGTGCTCTCTGCGAA 60
DB 589 ATGGCCCATGAGTGAATTGGAACCAATTGTACTGAGAGTGTGCTCTCTGCGAA 648
QY 61 AGTGAACGGAAGAGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
DB 649 AGTGAACGGAAGAGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 708
QY 121 CACATTTTGGAGCATTATATCACTGCTCCAGGTTTGAAGGAAAGTTGCAACAG 180
DB 709 CACATTTTGGAGCATTATATCACTGCTCCAGGTTTGAAGGAAAGTTGCAACAG 768
QY 181 GACAAAGTTTATTAACAGAGCTATCCAGCATTCAGCGAAGCATTAAGTTGACATTTAT 240
DB 769 GACAAAGTTTATTAACAGAGCTATCCAGCATTCAGCGAAGCATTAAGTTGACATTTAT 828
QY 241 TGCAGTCAGAGGTTGTTTACGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCTCA 300
DB 829 TGCAGTCAGAGGTTGTTTACGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCTCA 888
QY 301 GACTGTTTCTCAGTACTCTGGGTAACGTAAGAGGCTTCAACTGTTCACCTG 360
DB 889 GACTGTTTCTCAGTACTCTGGGTAACGTAAGAGGCTTCAACTGTTCACCTG 948
QY 361 CTTCGAGAGTGGTGTGCTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
DB 949 CTTCGAGAGTGGTGTGCTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 1008
QY 421 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCCCAACC 480
DB 1009 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCCCAACC 1068
QY 481 GGAATTTTCCCATCTTTATCTTGGCTGCAAGAGATGCTTCAACAGAGAGCTGATA 540
DB 1069 GGAATTTTCCCATCTTTATCTTGGCTGCAAGAGATGCTTCAACAGAGAGCTGATA 1128
QY 541 CAGCAGAGTGGATGTTGTTTAAATGCAAGCTATACTGTCCAAAGCTGACTTT 600
DB 1129 CAGCAGAGTGGATGTTGTTTAAATGCAAGCTATACTGTCCAAAGCTGACTTT 1188

QY 601 ATCCCGAGTCTCATTTTCCGCGTGGCCGTGATGACAGCTTTTGTGAGAAATTTTG 660
DB 1189 ATCCCGAGTCTCATTTTCCGCGTGGCCGTGATGACAGCTTTTGTGAGAAATTTTG 1248
QY 661 CCGTGTGTGACAAATCAGTGAATTTTCAATGAGAAAGCAAAAGCTTCAATGATGTT 720
DB 1249 CCGTGTGTGACAAATCAGTGAATTTTCAATGAGAAAGCAAAAGCTTCAATGATGTT 1308
QY 721 CTATGCTACTGTTTACTGAGATCTCCGCTCCGCCCAACCATCGCTTACATCATG 780
DB 1309 CTATGCTACTGTTTACTGAGATCTCCGCTCCGCCCAACCATCGCTTACATCATG 1368
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 840
DB 1369 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 1428
QY 841 ATATCTCCAAACTTCAATTTTCTGGCCCAACTCTGACTATGAGAAAGATTAAAGAC 900
DB 1429 ATATCTCCAAACTTCAATTTTCTGGCCCAACTCTGACTATGAGAAAGATTAAAGAC 1488
QY 901 CAGACT 906
DB 1489 CAGACT 1494
RESULT 12
ABV21080
ID ABV21080 standard; cDNA; 5145 BP.
XX
XX ABV21080;
AC
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 21071.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JR;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3481; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

Db 709 CACATTTGGAGCATTATATCACTGCTCCAGCTTATGAGGAGTTGCAACAG 768
Qy 181 GACAAAGTGTATATACAGAGCTCATCCAGCATTTCAAGGAAACATAAGTTGACATTGAT 240
Db 769 GACAAAGTGTATATACAGAGCTCATCCAGCATTTCAAGGAAACATAAGTTGACATTGAT 828
Qy 241 TGCAGTCAGAAAGTTGATTTAGATCAAGAGCTCCCAAGATGTTGGCTCTCTCTCTCA 300
Db 829 TGCAGTCAGAAAGTTGATTTAGATCAAGAGCTCCCAAGATGTTGGCTCTCTCTCTCA 888
Qy 301 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAAAGCTTCACTCTGTTCACTG 360
Db 889 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAAAGCTTCACTCTGTTCACTG 948
Qy 361 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
Db 949 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 1008
Qy 421 ACTCTAGTCCCTACCTGCAATTTCTCAGAGCTTGTGCTTACCTGTGCAACATTTGGGCCAACC 480
Db 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGAGCTTGTGCTTACCTGTGCAACATTTGGGCCAACC 1068
Qy 481 CGAATTTCTTCCCAATCTTTATCTTGTGCTGCGCAGAGAGATGCTTCAACAGAGCTGATA 540
Db 1069 CGAATTTCTTCCCAATCTTTATCTTGTGCTGCGCAGAGAGATGCTTCAACAGAGCTGATA 1128
Qy 541 CAGCAGAAATGGATGGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 600
Db 1129 CAGCAGAAATGGATGGTTATGTTTAAATGCCAGCAATCTGTCCAAAGCTGACTTT 1188
Qy 601 ATCCCGAGTCTCATTTCCGCGGTCCTGTGAATGACAGCTTTGTGAGAAATTTTG 660
Db 1189 ATCCCGAGTCTCATTTCCGCGGTCCTGTGAATGACAGCTTTGTGAGAAATTTTG 1248
Qy 661 CCGTGTGTGACAAATCATGATGATTTCTTGAAGAAAGCAAGCTTCAATGATGTT 720
Db 1249 CCGTGTGTGACAAATCATGATGATTTCTTGAAGAAAGCAAGCTTCAATGATGTT 1308
Qy 721 CTATGTCAGCTGTTTATGCTGGGATCTCCGCTCGGCACCAATCGCTATCGCTTACATATG 780
Db 1309 CTATGTCAGCTGTTTATGCTGGGATCTCCGCTCGGCACCAATCGCTTACATATG 1368
Qy 781 AAGAGGATGGACATGCTTATGATGAGCTTACAGATTTGGAAGAAAAAGCTACT 840
Db 1369 AAGAGGATGGACATGCTTATGATGAGCTTACAGATTTGGAAGAAAAAGCTACT 1428
Qy 841 ATATCTCCAAACTTCAATTTCTGGGCAACTCCTGAGCTATGAGAGAAATTAAGAAC 900
Db 1429 ATATCTCCAAACTTCAATTTCTGGGCAACTCCTGAGCTATGAGAGAAATTAAGAAC 1488
Qy 901 CAGACT 906
Db 1489 CAGACT 1494

RESULT 14
ABV20978
ID ABV20978 standard; cDNA; 5145 BP.
XX AC ABV20978;
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 20969.
XX KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KM pharmacogenomic marker; gene; ss.
OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-021314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX DR Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 3451; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Query Match 99.6%; Score 902.8; DB 5; Length 5145;
Best Local Similarity 99.8%; Pred. No. 8e-278; 2; Indels 0; Gaps 0;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCCCATGATGATGATGGAATGGAATGTTTACTGAGAGTGTGCTGCTGGAA 60
Db 589 ATGGCCCATGATGATGGAATGGAATGTTTACTGAGAGTGTGCTGCTGGAA 648
Qy 61 AGTGAACGGAAGAGTCTGCTTAATGATAGCCGCAATTTGTGAAATCATATCC 120
Db 649 AGTGAACGGAAGAGTCTGCTTAATGATAGCCGCAATTTGTGAAATCATATCC 708
Qy 121 CAAATTTGGAACCTTAATATCACTGCTCAAGCTTATGAAAGGAAGTTGCAACAG 180
Db 709 CAAATTTGGAACCTTAATATCACTGCTCAAGCTTATGAAAGGAAGTTGCAACAG 768
Qy 181 GACAAAGTGTATATACAGAGCTCATCCAGCATTTCAAGGAAACATAAGTTGACATTGAT 240
Db 769 GACAAAGTGTATATACAGAGCTCATCCAGCATTTCAAGGAAACATAAGTTGACATTGAT 828
Qy 241 TGCAGTCAGAAAGTTGATTTAGATCAAGAGCTCCCAAGATGTTGGCTCTCTCTCTCA 300
Db 829 TGCAGTCAGAAAGTTGATTTAGATCAAGAGCTCCCAAGATGTTGGCTCTCTCTCTCA 888
Qy 301 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAAAGCTTCACTCTGTTCACTG 360
Db 889 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAAAGCTTCACTCTGTTCACTG 948
Qy 361 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
Db 949 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 1008
Qy 421 ACTCTAGTCCCTACCTGCAATTTCTCAGAGCTTGTGCTTACCTGTGCAACATTTGGGCCAACC 480
Db 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGAGCTTGTGCTTACCTGTGCAACATTTGGGCCAACC 1068

QY 481 CGAATCTCCCAATCTTATCTTGGCTGCCAGGAGATGTCTCAACAAGAGCTGATA 540
 DB 1069 CGAATCTCTCCCAATCTTATCTTGGCTGCCAGGAGATGTCTCAACAAGAGCTGATA 1128
 QY 541 CAGCAGATGGAGTTGGTTATGTGTTAAATGCGAGTATCTCTCAAGCTGACTTT 600
 DB 1129 CAGCAGAAATGGGATGGTTATGTGTTAAATGCGAGTATCTCTCAAGCTGACTTT 1188
 QY 601 ATCCCGAGTCTCAATTCCTCCGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 660
 DB 1189 ATCCCGAGTCTCAATTCCTCCGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 1248
 QY 661 CCGTGTGGTGAACAATATGATGATTCATTTGAGAAAGCAAAAGCTCCCAATGAGTGTG 720
 DB 1249 CCGTGTGGTGAACAATATGATGATTCATTTGAGAAAGCAAAAGCTCCCAATGAGTGTG 1308
 QY 721 CTAGTGCACTGTTTAAAGTGGGATCTCCGCTCGGCCACCATGCTATCGCTACATCATG 780
 DB 1309 CTAGTGCACTGTTTAAAGTGGGATCTCCGCTCGGCCACCATGCTATCGCTACATCATG 1368
 QY 781 AAGAGATGACATGCTTTTAAATGATGAGCTTACAGATTTGTGAAAGAAAGCTTACT 840
 DB 1369 AAGAGATGACATGCTTTTAAATGATGAGCTTACAGATTTGTGAAAGAAAGCTTACT 1428
 QY 841 ATATCTCAAACTTCAATTTCTGGGCCCACTCCCTGACCTATGAGAAAGATTTAGAAC 900
 DB 1429 ATATCTCAAACTTCAATTTCTGGGCCCACTCCCTGACCTATGAGAAAGATTTAGAAC 1488
 QY 901 CAGACT 906
 DB 1489 CAGACT 1494

RESULT 15
 ABV21092
 ID ABV21092 standard; cDNA; 5145 BP.

AC ABV21092;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 21083.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; se.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Endege WO, Monahan JR;
 XX
 DR MPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PS for detecting presence of prostate cancer, stage of prostate cancer.
 XX Claim 1; Page 3485; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
 Query Match 99.6%; Score 902.8; DB 5; Length 5145;
 Best Local Similarity 99.8%; Pred. No. 8e-278;
 Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGGCCCATGATGATGGAAGTCAATTTGTACTGAGAGGTGTGCTCTGCTGAA 60
 DB 589 ATGGCCCATGATGATGGAAGTCAATTTGTACTGAGAGGTGTGCTCTGCTGAA 648
 QY 61 AGTGAACGGAAGAAAGTGTGCTGCTAATTTGATGAGCCGCAATTTGTGAAATACATATCC 120
 DB 649 AGTGAACGGAAGAAAGTGTGCTGCTAATTTGATGAGCCGCAATTTGTGAAATACATATCC 708
 QY 121 CACATTTGGAAAGCTTAAATATCAATGCTCCAGCTTATGAAAGCAAGTTGCAACAG 180
 DB 709 CACATTTGGAAAGCTTAAATATCAATGCTCCAGCTTATGAAAGCAAGTTGCAACAG 768
 QY 181 GACAAAGTGTAAATACAGAGCTCATCCAGATTCAGCGAAACCTAAAGTTGACATTAAT 240
 DB 769 GACAAAGTGTAAATACAGAGCTCATCCAGATTCAGCGAAACCTAAAGTTGACATTAAT 828
 QY 241 TCGAGTCGGAAGGTGTGTTTAAATGATCAAGTCCCAAGATGTGCTCTCTCTTCA 300
 DB 829 TCGAGTCGGAAGGTGTGTTTAAATGATCAAGTCCCAAGATGTGCTCTCTCTTCA 888
 QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACGTGAGAGGCTTCACTCTGTCACTG 360
 DB 889 GACTGTTTCTCACTGATCTTCTGGGTAACGTGAGAGGCTTCACTCTGTCACTG 948
 QY 361 CTTCGAGGTGGTGTGAGTCTCTGTTTCCCTGGCTCTGTGAAAGAAATATCC 420
 DB 949 CTTCGAGGTGGTGTGAGTCTCTGTTTCCCTGGCTCTGTGAAAGAAATATCC 1008
 QY 421 ACTTAAGTCCCTAATGCAATTTCTCAGCTTGTACCGTTGCAAAATGAGGCAACC 480
 DB 1009 ACTTAAGTCCCTAATGCAATTTCTCAGCTTGTACCGTTGCAAAATGAGGCAACC 1068
 QY 481 CGAATCTCCCAATCTTATCTTGGCTGCCAGGAGATGTCTCAACAAGAGCTGATA 540
 DB 1069 CGAATCTCTCCCAATCTTATCTTGGCTGCCAGGAGATGTCTCAACAAGAGCTGATA 1128
 QY 541 CAGCAGATGGAGTTGGTTATGTGTTAAATGCGAGTATCTCTCAAGCTGACTTT 600
 DB 1129 CAGCAGAAATGGGATGGTTATGTGTTAAATGCGAGTATCTCTCAAGCTGACTTT 1188
 QY 601 ATCCCGAGTCTCAATTCCTCCGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 660
 DB 1189 ATCCCGAGTCTCAATTCCTCCGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 1248
 QY 661 CCGTGTGGTGAACAATATGATGATTCATTTGAGAAAGCAAAAGCTCCCAATGAGTGTG 720
 DB 1249 CCGTGTGGTGAACAATATGATGATTCATTTGAGAAAGCAAAAGCTCCCAATGAGTGTG 1308
 QY 721 CTAGTGCACTGTTTAAAGTGGGATCTCCGCTCGGCCACCATGCTATCGCTACATCATG 780
 DB 1309 CTAGTGCACTGTTTAAAGTGGGATCTCCGCTCGGCCACCATGCTATCGCTACATCATG 1368
 QY 781 AAGAGATGACATGCTTTTAAATGATGAGCTTACAGATTTGTGAAAGAAAGCTTACT 840

Db	1369	AAGAGATGACATGCTTTAGATGAAGCTTACGATTGTGAAGAAAAAGACCTACT	1428
Qy	841	ATATCTCCAAACTTTCATTTTCTGGCCACTCTGACTATGAGAAGATTAGAAC	900
Db	1429	ATATCTCCAAACTTTCATTTTCTGGCCACTCTGACTATGAGAAGATTAGAAC	1488
Qy	901	CAGACT	906
Db	1489	CAGACT	1494

Search completed: June 21, 2004, 18:35:46
Job time : 528.548 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 18:08:06 ; Search time 540.914 Seconds
(without alignments)
7673.060 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 906
Sequence: 1 atggcccaatgagatgattg9.....agaagattaagaccagact 906

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	902.8	99.6	1998	6	US-09-816-494-3
2	902.8	99.6	1998	16	US-10-377-072-27
3	902.8	99.6	2071	13	US-10-072-012-257
4	902.8	99.6	2732	13	US-10-168-506-2
5	902.8	99.6	3059	17	US-10-257-026-1
6	902.8	99.6	3496	9	US-09-964-277-1
7	902.8	99.6	3544	9	US-09-816-494-1
8	902.8	99.6	3544	16	US-10-377-072-25
9	902.8	99.6	3625	13	US-10-425-114-26234
10	902.8	99.6	3766	13	US-10-343-357-17
11	902.8	99.6	4790	17	US-10-648-593-115
12	901.2	99.5	2102	16	US-10-094-749-673
13	899.6	99.3	2966	13	US-10-296-115-520
14	817.8	90.3	2200	13	US-10-072-012-255

15	564.8	62.3	3332	9	US-09-964-277-20	Sequence 20, Appl
16	363.6	40.1	2453	14	US-10-005-858-1	Sequence 1, Appl
17	335.8	37.1	2476	13	US-10-220-120-28	Sequence 28, Appl
18	290.8	32.1	2039	13	US-10-072-012-265	Sequence 265, Appl
19	272.4	30.1	877	9	US-09-764-853-158	Sequence 158, Appl
20	223	24.6	277	17	US-10-637-855-29	Sequence 29, Appl
21	223	24.6	279	16	US-10-305-720-91	Sequence 91, Appl
22	217	24.0	246	16	US-10-257-026-3	Sequence 3, Appl
23	124.6	13.8	1337	9	US-09-964-899-46	Sequence 46, Appl
24	124.6	13.8	1449	12	US-10-052-482-192	Sequence 192, Appl
25	124.6	13.8	1635	17	US-10-648-593-27	Sequence 27, Appl
26	124.6	13.8	1830	15	US-10-346-356-1	Sequence 1, Appl
27	124.6	13.8	1909	12	US-10-052-482-191	Sequence 191, Appl
28	124.6	13.8	1909	13	US-10-058-270A-129	Sequence 129, Appl
29	124.6	13.8	3286	10	US-09-971-329-9	Sequence 9, Appl
30	115	12.7	1452	12	US-10-052-482-189	Sequence 189, Appl
31	115	12.7	1977	12	US-10-052-482-188	Sequence 188, Appl
32	110.6	12.2	1356	15	US-10-184-832-6	Sequence 6, Appl
33	110.6	12.2	2741	15	US-10-184-832-4	Sequence 4, Appl
34	106.8	11.8	2303	9	US-09-880-107-3849	Sequence 3849, Ap
35	106.8	11.8	2303	10	US-09-922-146-3	Sequence 3, Appl
36	106.8	11.8	2303	13	US-10-342-887-603	Sequence 603, Appl
37	106.8	11.8	2303	13	US-10-172-118-603	Sequence 603, Appl
38	106.8	11.8	2338	13	US-10-363-616-7	Sequence 7, Appl
39	106.6	11.8	1143	15	US-10-184-832-3	Sequence 3, Appl
40	106.6	11.8	1574	13	US-10-296-115-440	Sequence 440, Appl
41	106.6	11.8	2650	15	US-10-184-832-1	Sequence 1, Appl
42	103.4	11.4	2104	16	US-10-152-312A-2041	Sequence 2041, Ap
43	103.4	11.4	2104	16	US-10-388-934-654	Sequence 654, Appl
44	103.4	11.4	2104	16	US-10-191-803-270	Sequence 270, Appl
45	99.2	10.9	462	10	US-09-918-995-31436	Sequence 31436, A

ALIGNMENTS

RESULT 1
US-09-816-494-3
Sequence 3, Application US/09816494
Patent No. US20020034807A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OR INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
TITLE OR INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1998
TYPE: DNA
ORGANISM: Homo sapiens
US-09-816-494-3

Query Match	99.6%	Score	902.8	DB	9	Length	1998
Best Local Similarity	99.6%	Pred. No.	1,4e-282				
Matches	904	Conservative	0	Mismatches	2	Indels	0
						Gaps	0
QY	1	ATGGCCCATGAGTGAATGTAAGTGAAGGTTGGTGGCTGCTGGAA	60				
DB	1	ATGGCCCATGAGTGAATGTAAGTGAAGGTTGGTGGCTGCTGGAA	60				
QY	61	AGTGAAGCGAAGAGTGTCTGTAATGATGAGCGGCATTTGTGATCAATACATCC	120				
DB	61	AGTGAAGCGAAGAGTGTCTGTAATGATGAGCGGCATTTGTGATCAATACATCC	120				
QY	121	CACATTTGGAAGCCATTAATCACTGCTCAAGCTTATGAAGCAAGTTGCAACAG	180				
DB	121	CACATTTGGAAGCCATTAATCACTGCTCAAGCTTATGAAGCAAGTTGCAACAG	180				

```
QY 181 GACAAAGTGTAAATTCAGAGCTCATCCAGCTTCAAGGAAACATAAGTTGACATTGAT 240
Db 181 GACAAAGTGTAAATTCAGAGCTCATCCAGCTTCAAGGAAACATAAGTTGACATTGAT 240
QY 241 TGCAGTCAGAAAGTGTAGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 300
Db 241 TGCAGTCAGAAAGTGTAGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 300
QY 301 GACTGTTTTCTCACTGTAATCTCTGGGTAAACTGGAAGAGCTTCAACTCTGTTCACTG 360
Db 301 GACTGTTTTCTCACTGTAATCTCTGGGTAAACTGGAAGAGCTTCAACTCTGTTCACTG 360
QY 361 CTGAGAGTGGGTTTCTGAGTTCTCTGTTGTTCCCTGGGCTCTGTAAGAAATCC 420
Db 361 CTGAGAGTGGGTTTCTGAGTTCTCTGTTGTTCCCTGGGCTCTGTAAGAAATCC 420
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCCCTGCTTAACCTGTTGCCAAATTGGGCAACC 480
Db 421 ACTCTAGTCCCTACCTGATTTCTCAGCCCTGCTTAACCTGTTGCCAAATTGGGCAACC 480
QY 481 CGAATTTCTCCCAATCTTTATCTTGCTGCTGCAAGAGATGCTTCAAGAGAGCTGATA 540
Db 481 CGAATTTCTCCCAATCTTTATCTTGCTGCTGCAAGAGATGCTTCAAGAGAGCTGATA 540
QY 541 CAGCAGAAATGGAGTTGTTATGTTAAATGCACTTACCTGTCAAAGCCCTGACTTT 600
Db 541 CAGCAGAAATGGAGTTGTTATGTTAAATGCACTTACCTGTCAAAGCCCTGACTTT 600
QY 601 ATCCCGAGTCTATTTCTGCTGTGCTGCTGATGATGACAGCTTTGTTGAGAAATTTTG 660
Db 601 ATCCCGAGTCTATTTCTGCTGTGCTGCTGATGATGACAGCTTTGTTGAGAAATTTTG 660
QY 661 CCGTGGTTGGAACAAATCAGATGATTTCTAGAGAAACAAAGCTTCCAAATGATGTTG 720
Db 661 CCGTGGTTGGAACAAATCAGATGATTTCTAGAGAAACAAAGCTTCCAAATGATGTTG 720
QY 721 CTAGTGCACCTGTTTACGTTGAGATCTCCGCTCCGCCACATGCTATCCCTACATCAG 780
Db 721 CTAGTGCACCTGTTTACGTTGAGATCTCCGCTCCGCCACATGCTATCCCTACATCAG 780
QY 781 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAGAGCTTACT 840
Db 781 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAGAGCTTACT 840
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTCGAGCTATGAGAAGAGTTTAAAGAC 900
Db 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTCGAGCTATGAGAAGAGTTTAAAGAC 900
QY 901 CAGACT 906
Db 901 CAGACT 906
```

RESULT 2

```
US-10-377-072-27
; Sequence 27, Application US/10377072
; Publication No. US2004009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: MP103-0180NM1M
; CURRENT APPLICATION NUMBER: US/10/377, 072
; CURRENT FILING DATE: 2003-02-27
```

```

; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)
US-10-377-072-27
```

```
Query Match 99.6%; Score 902.8; DB 16; Length 1998;
Best Local Similarity 99.8%; Pred. No. 1.4e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCCCATGAGATGATTTGAACTCAATTTGTTACTGAGAGTTGGTGGCTCTGCTGGA 60
Db 1 ATGGGCCCATGAGATGATTTGAACTCAATTTGTTACTGAGAGTTGGTGGCTCTGCTGGA 60
QY 61 AGTGAACGGAAGAAAGTCTCTAATTGATGAGCGGCAATTTGTAATPACATATCCTC 120
Db 61 AGTGAACGGAAGAAAGTCTCTAATTGATGAGCGGCAATTTGTAATPACATATCCTC 120
QY 121 CACATTTTGAAGCATTAATATCACTGCTCCAGCTTATGAGGAAGTTGGAAG 180
Db 121 CACATTTTGAAGCATTAATATCACTGCTCCAGCTTATGAGGAAGTTGGAAG 180
QY 181 GACAAAGTGTAAATTCAGAGCTCATCCAGCTTCAAGGAAACATAAGTTGACATTGAT 240
Db 181 GACAAAGTGTAAATTCAGAGCTCATCCAGCTTCAAGGAAACATAAGTTGACATTGAT 240
QY 241 TGCAGTCAGAAAGTGTAGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 300
Db 241 TGCAGTCAGAAAGTGTAGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 300
QY 301 GACTGTTTTCTCACTGTAATCTCTGGGTAAACTGGAAGAGCTTCAACTCTGTTCACTG 360
Db 301 GACTGTTTTCTCACTGTAATCTCTGGGTAAACTGGAAGAGCTTCAACTCTGTTCACTG 360
QY 361 CTGAGAGTGGGTTTCTGAGTTCTCTGTTGTTCCCTGGGCTCTGTAAGAAATCC 420
Db 361 CTGAGAGTGGGTTTCTGAGTTCTCTGTTGTTCCCTGGGCTCTGTAAGAAATCC 420
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCCCTGCTTAACCTGTTGCCAAATTGGGCAACC 480
Db 421 ACTCTAGTCCCTACCTGATTTCTCAGCCCTGCTTAACCTGTTGCCAAATTGGGCAACC 480
QY 481 CGAATTTCTCCCAATCTTTATCTTGCTGCTGCAAGAGATGCTTCAAGAGAGCTGATA 540
Db 481 CGAATTTCTCCCAATCTTTATCTTGCTGCTGCAAGAGATGCTTCAAGAGAGCTGATA 540
QY 541 CAGCAGAAATGGAGTTGTTATGTTAAATGCACTTACCTGTCAAAGCCCTGACTTT 600
```


Db	541	CAGCAGAAATGGGATTCGTTATGTTAAATGCAGCAAAATCTGTCCAAAGCTGACTTT	6000
Qy	601	ATCCCCGAGTCATCTTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG	6606
Db	601	ATCCCCGAGTCATCTTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG	6606
Qy	661	CCGTGGTGGACAAATCAGTAGATTTCAATTGAGAAAGCAAAAGCCTCCAAATGATGTGT	7200
Db	661	CCGTGGTGGACAAATCAGTAGATTTCAATTGAGAAAGCAAAAGCCTCCAAATGATGTGT	7200
Qy	721	CTAGTGCACCTGTTTAGCTGGGATCTCCCGCTCCGCCACCAATCGCTATTCGCTTACATCATG	7800
Db	721	CTAGTGCACCTGTTTAGCTGGGATCTCCCGCTCCGCCACCAATCGCTTACATCATG	7800
Qy	781	AAGAGGATGAGCATGCTTTAATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT	8400
Db	781	AAGAGGATGAGCATGCTTTAATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT	8400
Qy	841	ATATCTCCAAACTTCATTTTCTGGGCCAACTCTGACATATGAGAGAAAGATTTAAGAAC	9000
Db	841	ATATCTCCAAACTTCATTTTCTGGGCCAACTCTGACATATGAGAGAAAGATTTAAGAAC	9000
Qy	901	CAGACT 906	
Db	901	CAGACT 906	

RESULT 3

```

US-10-072-012-257
/ Sequence 257, Application US/10072012
/ Publication No. US2004003493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Spyrek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Patnuretan, Meera
/ APPLICANT: Shimketa, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangoli, Esha
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Colman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grosse, William M.
/ APPLICANT: Alsebrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072,012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767

```

Query Match	99.6%	Score 902.8	DB 13	Length 2071
Best Local Similarity	99.8%	Pred. No. 1.5e-282		
Matches 904	Conservative 0	Mismatches 2	Indels 0	Gaps 0
US-10-072-012-257				
ORGANISM: Homo sapiens				
SEQ ID NO 257				
LENGTH: 2071				
TYPE: DNA				
Query Match	99.6%	Score 902.8	DB 13	Length 2071
Best Local Similarity	99.8%	Pred. No. 1.5e-282		
Matches 904	Conservative 0	Mismatches 2	Indels 0	Gaps 0
1	ATGGCCCATGAGATGATTGGAACTCAATTGTACTGAGAGGTGGTGGCTCTGCTGGAA	60		
61	ATGGCCCATGAGATGATTGGAACTCAATTGTACTGAGAGGTGGTGGCTCTGCTGGAA	120		
61	AGTGAAGAGAAAAAGTGTCTGTAATTATATGACGGGACATTGTGTGAATACATATCC	120		
121	AGTGAAGAGAAAAAGTGTCTGTAATTATATGACGGGACATTGTGTGAATACATATCC	180		
121	CACATTTTGAAGCATTATATATCACTGCTCCAACTTATATGAGCGAAGTTGCAACAG	180		
181	CACATTTTGAAGCATTATATATCACTGCTCCAACTTATATGAGCGAAGTTGCAACAG	240		
181	GACAAAGTGTATATTAACAGACCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT	240		
241	GACAAAGTGTATATTAACAGACCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT	300		
241	GACAAAGTGTATATTAACAGACCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT	360		
301	TGCAGTCAGAAAGTGTATGTTATGATCAAAAGCTCCAAAGATGTGCTCTCTCTTCA	360		
301	GACTGTTTCTCATCTGTACTCTGAGGTAACTGAGAAAGAGCTTCACTGTTCACCTG	360		
361	GACTGTTTCTCATCTGTACTCTGAGGTAACTGAGAAAGAGCTTCACTGTTCACCTG	420		
361	CTTGACAGGTGGTGTGCTGAGTCTCTGCTGTTTCCCTGCGCTCTGTGAAGAAAAATCC	480		
421	CTTGACAGGTGGTGTGCTGAGTCTCTGCTGTTTCCCTGCGCTCTGTGAAGAAAAATCC	480		
421	ACTCTAGTCTCTACTCTGATTTCTCAGGCTTGCTTACTCTGTTCCAACTTGGGCAACC	480		
481	ACTCTAGTCTCTACTCTGATTTCTCAGGCTTGCTTACTCTGTTCCAACTTGGGCAACC	540		
481	GGAATTTCTCCCAATCTTTATCTTGAGGCGCAGAGAGATGTCTCAACAAAGAGCTGATA	540		
541	GGAATTTCTCCCAATCTTTATCTTGAGGCGCAGAGAGATGTCTCAACAAAGAGCTGATA	600		
541	CAGCAGATGAGGATGTGTTATGTGTTAAATGCAAGCAATTAATCTGTCCAAAGCTGACCTT	660		
601	CAGCAGATGAGGATGTGTTATGTGTTAAATGCAAGCAATTAATCTGTCCAAAGCTGACCTT	660		
601	ATCCCGAGTCTCATTTCTCTGCGGTGCTGTGATGACAGCTTTTGTGAAGAAAAATTTTG	660		
661	ATCCCGAGTCTCATTTCTCTGCGGTGCTGTGATGACAGCTTTTGTGAAGAAAAATTTTG	720		
661	CCGTGTTTGAACAATCAGTATGATTTCTATGAGAAAGCAAAAGCTCCAAATGAGATGTGT	720		
721	CCGTGTTTGAACAATCAGTATGATTTCTATGAGAAAGCAAAAGCTCCAAATGAGATGTGT	780		
721	CTAGTGCACTGTTAGCTGGAGCTCCGCTCCGCAACATTCGCTATGCGCTCAATCATG	780		
781	CTAGTGCACTGTTAGCTGGAGCTCCGCTCCGCAACATTCGCTATGCGCTCAATCATG	840		
781	AAGAGGATGACATGCTCTTTTGAATGAGACTTACAGATTTGTGAAGAAAAAGACTACT	840		

Db 841 AAGAGATGACATGCTTACATGAGCTTACAAATTGGAAGAAAAAGACTACT 900
QY 841 ATATCTCCAACTTCAATTTCTGCGCCAACTCTGAGCTATGAGAGAGATTAGAAC 900
Db 901 ATATCTCCAACTTCAATTTCTGCGCCAACTCTGAGCTATGAGAGAGATTAGAAC 900
QY 901 CAGACT 906
Db 961 CAGACT 966

RESULT 4

US-10-168-506-2
; Sequence 2, Application US/10168506
; Publication No. US20040053229A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERRARD
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: HILL, RON
; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
; FILE REFERENCE: 038602/1351
; CURRENT APPLICATION NUMBER: US/10/168,506
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: PCT/US00/34736
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-168-506-2

Query Match 99.6%; Score 902.8; DB 13; Length 2732;
Best Local Similarity 99.8%; Pred. No. 1.7e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCCCATGAGATGATTTGGAACCTCAATTGTTACTGAGAGGTGGTCTGCTGGA 60
Db 538 ATGCCCATGAGATGATTTGGAACCTCAATTGTTACTGAGAGGTGGTCTGCTGGA 60
QY 61 AGTGAACGGAAGAAAGTCTGCTTAATTGATGAGCGGCAATTGTTGGAATCAATATCC 120
Db 598 AGTGAACGGAAGAAAGTCTGCTTAATTGATGAGCGGCAATTGTTGGAATCAATATCC 120
QY 121 CACATTTTGAAGCCATTAATATCAAGCTCCCAAGCTTATGAAAGCAAGTTGCAACG 657
Db 658 CACATTTTGAAGCCATTAATATCAAGCTCCCAAGCTTATGAAAGCAAGTTGCAACG 657
QY 181 GACAAAGTTTAATACAGAGCTATCCAGCATTCAGGAAACATAAGTTGATGAT 240
Db 718 GACAAAGTTTAATACAGAGCTATCCAGCATTCAGGAAACATAAGTTGATGAT 240
QY 241 TGCAGTCAAGAGTTGATGATTAAGTCAAAAGCTCCCAAGTTGCTCTCTCTTCA 300
Db 778 TGCAGTCAAGAGTTGATGATTAAGTCAAAAGCTCCCAAGTTGCTCTCTCTTCA 300
QY 301 GACTGTTTCTCACTGTAATCTTGGTAACTGGAAGAGCTTCAACTGTGCACTG 360
Db 838 GACTGTTTCTCACTGTAATCTTGGTAACTGGAAGAGCTTCAACTGTGCACTG 360
QY 361 CTTCAGAGTGGTGTGCTGAGTCTCTCGTGTGTTCCCTGCGCTCTGGAAGAAATCC 420
Db 898 CTTCAGAGTGGTGTGCTGAGTCTCTCGTGTGTTCCCTGCGCTCTGGAAGAAATCC 420
QY 421 ACTTAGTCCCTACCTGCAATTTCTCAGCTTCTTACCTGTTGCAACATTTGGGCAAC 480
Db 958 ACTTAGTCCCTACCTGCAATTTCTCAGCTTCTTACCTGTTGCAACATTTGGGCAAC 480

QY 481 CGAATTTCTCCCATCTTTATCTTGGCTGCGCAGCAGATGCTTCAACAAGSAGCTGATA 540
Db 1018 CGAATTTCTCCCATCTTTATCTTGGCTGCGCAGCAGATGCTTCAACAAGSAGCTGATA 540
QY 541 CAGCAGAAATGGAGTTGTTATGTTAAATCCAGCTTATCTCTCCAAAGCTGACTTT 600
Db 1078 CAGCAGAAATGGAGTTGTTATGTTAAATCCAGCTTATCTCTCCAAAGCTGACTTT 600
QY 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
Db 1138 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
QY 661 CCGTGGTGGACAAATCAGTATGATTTCAATTGAGAAAGCAAAAGCTCCATGAGTGT 720
Db 1198 CCGTGGTGGACAAATCAGTATGATTTCAATTGAGAAAGCAAAAGCTCCATGAGTGT 720
QY 721 CTAGTGCACTGTTTACGCTGGGATCTCCCGCTCCGCAACATGCTATGCTTACATCAT 780
Db 1258 CTAGTGCACTGTTTACGCTGGGATCTCCCGCTCCGCAACATGCTTACATCATCAT 780
QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGTTTGTGAAGAAAAAGACTACT 840
Db 1318 AAGAGATGACATGCTTTAGATGAGCTTACAGTTTGTGAAGAAAAAGACTACT 840
QY 841 ATATCTCCAACTTCAATTTCTGCGCCAACTCTGAGCTATGAGAGAGATTAGAAC 900
Db 1378 ATATCTCCAACTTCAATTTCTGCGCCAACTCTGAGCTATGAGAGAGATTAGAAC 900
QY 901 CAGACT 906
Db 1438 CAGACT 1443

RESULT 5

US-10-257-026-1
; Sequence 1, Application US/10257026
; Publication No. US20040086859A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10KDS
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3059
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (127)...(2121)
US-10-257-026-1

Query Match 99.6%; Score 902.8; DB 17; Length 3059;
Best Local Similarity 99.8%; Pred. No. 1.9e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCCCATGAGATGATTTGGAACCTCAATTGTTACTGAGAGGTGGTCTGCTGGA 60
Db 127 ATGCCCATGAGATGATTTGGAACCTCAATTGTTACTGAGAGGTGGTCTGCTGGA 60
QY 61 AGTGAACGGAAGAAAGTCTGCTTAATTGATGAGCGGCAATTGTTGGAATCAATATCC 120
Db 187 AGTGAACGGAAGAAAGTCTGCTTAATTGATGAGCGGCAATTGTTGGAATCAATATCC 120
QY 121 CACATTTTGAAGCCATTAATCACTGCTCAAGCTTATGAGAGAGAGTTGCAACG 246
Db 247 CACATTTTGAAGCCATTAATCACTGCTCAAGCTTATGAGAGAGAGTTGCAACG 246
QY 181 GACAAAGTTTAATACAGAGCTATCCAGCATTCAGGAAACATAAGTTGATGAT 240
Db 181 GACAAAGTTTAATACAGAGCTATCCAGCATTCAGGAAACATAAGTTGATGAT 240

Db 307 GACAAAGTGTAAATACAGAGCTCATCCAGCAATTCAGCGAAACATAGTTGACATGAT 366
Qy 241 TGCAGTCAGAAAGTTGATGTTAGATCAAGATCCCAAGATGTTGGCTCTCTCTTCA 300
Db 367 TGCAGTCAGAAAGTTGATGTTAGATCAAGATCCCAAGATGTTGGCTCTCTCTTCA 426
Qy 301 GACGTTTTCTCACTGACTTCTGGGTAACTGGAGAAAGCTTCAACTGTTACCTG 360
Db 427 GACTGTTTTCTCACTGACTTCTGGGTAACTGGAGAAAGCTTCAACTGTTACCTG 486
Qy 361 CTTCAGAGTGGTGTGCTGAGATCTCTGTTGTTCCCTGGCTCTGCTGAGAAATCC 420
Db 487 CTTCAGAGTGGTGTGCTGAGATCTCTGTTGTTCCCTGGCTCTGCTGAGAAATCC 546
Qy 421 ACTCTAGTCCCTTACCTGATCTTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 547 ACTCTAGTCCCTTACCTGATCTTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTG 606
Qy 481 CGAATCTTCCCAATCTTATCTTGGCTGCGAGAGATGCTCCCAAGAGCTGATA 540
Db 607 CGAATCTTCCCAATCTTATCTTGGCTGCGAGAGATGCTCCCAAGAGAGCTGATG 666
Qy 541 CACAGATGGAGATGGTATGTTAAATGCAAGCTATACCTGCTCAAGCTGACTTT 600
Db 667 CACAGATGGAGATGGTATGTTAAATGCAAGCTATACCTGCTCAAGCTGACTTT 726
Qy 601 ATCCCGAGTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 727 ATCCCGAGTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
Qy 661 CGGTGGTGGCAAAATAGTATGATTTGATGAGAAAGCAAAAGCTCCCAATGATGTT 720
Db 787 CGGTGGTGGCAAAATAGTATGATTTGATGAGAAAGCAAAAGCTCCCAATGATGTT 846
Qy 721 CTAGTGACTGTTAGTGGGATCTCCGCTCGGCAACATGCTATCGCTACATCATG 780
Db 847 CTAGTGACTGTTAGTGGGATCTCCGCTCGGCAACATGCTATCGCTACATCATG 906
Qy 781 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAGAAAAAGCTTACT 840
Db 907 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAGAAAAAGCTTACT 966
Qy 841 ATATCTCCAAACTTCAATTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
Db 967 ATATCTCCAAACTTCAATTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 1026
Qy 901 CAGACT 906
Db 1027 CAGACT 1032

RESULT 6

US-09-964-277-1
; Sequence 1, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-1

Query Match 99.6%; Score 902.8; DB 9; Length 3496;
Best Local Similarity 99.8%; Pred. No. 2e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATGGGAATCAAAATTTGTTACTGAGAGTTGGGCTCTGCTGAA 60
Db 562 ATGGCCCATGAGATGATGGGAATCAAAATTTGTTACTGAGAGTTGGGCTCTGCTGAA 621
Qy 61 AGTGAACGGAAGAAAGTGTGCTAATGATAGCGGCAATTTGTGAAATCAATACATCC 120
Db 622 AGTGAACGGAAGAAAGTGTGCTAATGATAGCGGCAATTTGTGAAATCAATACATCC 681
Qy 121 CACATTTGGAGCCATTAATATCAATGCTCCAGCTTATGAAAGGAGTTGCAACAG 180
Db 682 CACATTTGGAGCCATTAATATCAATGCTCCAGCTTATGAAAGGAGTTGCAACAG 741
Qy 181 GACAAAGTGTAAATTAAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTTGAT 240
Db 742 GACAAAGTGTAAATTAAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTTGAT 801
Qy 241 TGCAGTCAGAAAGTTGATGTTACGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 300
Db 802 TGCAGTCAGAAAGTTGATGTTACGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 861
Qy 301 GACTGTTTTCTCACTGACTTCTGGGTAACTGGAGAGCTTCAACTGTTTCACTG 360
Db 862 GACTGTTTTCTCACTGACTTCTGGGTAACTGGAGAGCTTCAACTGTTTCACTG 921
Qy 361 CTTCAGAGTGGTGTGCTGAGATCTCTGTTGTTCCCTGGCTCTGCTGAGAAATCC 420
Db 922 CTTCAGAGTGGTGTGCTGAGATCTCTGTTGTTCCCTGGCTCTGCTGAGAAATCC 981
Qy 421 ACTCTAGTCCCTTACCTGATCTTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 982 ACTCTAGTCCCTTACCTGATCTTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1041
Qy 481 CGAATCTTCCCAATCTTATCTTGGCTGCGAGAGATGCTCTCAAGAGCTGATA 540
Db 1042 CGAATCTTCCCAATCTTATCTTGGCTGCGAGAGATGCTCTCAAGAGAGCTGATG 1101
Qy 541 CACAGAAATGGAGATGGTATGTTAAATGCAAGCTATACCTGCTCAAGAGCTGACTT 600
Db 1102 CACAGAAATGGAGATGGTATGTTAAATGCAAGCTATACCTGCTCAAGAGCTGACTT 1161
Qy 601 ATCCCGAGTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 1162 ATCCCGAGTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1221
Qy 661 CGGTGGTGGCAAAATGATGATTTCAATTTGAGAAAGCAAAAGCTTCAATGATGTT 720
Db 1222 CGGTGGTGGCAAAATGATGATTTCAATTTGAGAAAGCAAAAGCTTCAATGATGTT 1281
Qy 721 CTAGTGACTGTTAGTGGGATCTCCGCTCGGCAACATGCTATCGCTACATCATG 780
Db 1282 CTAGTGACTGTTAGTGGGATCTCCGCTCGGCAACATGCTATCGCTACATCATG 1341
Qy 781 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAGAAAAAGCTTACT 840
Db 1342 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAGAAAAAGCTTACT 1401
Qy 841 ATATCTCCAAACTTCAATTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
Db 1402 ATATCTCCAAACTTCAATTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 1461
Qy 901 CAGACT 906
Db 1462 CAGACT 1467

RESULT 7

US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY

```

; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)...(2583)
US-09-816-494-1

```

```

Query Match      99.6%; Score 902.8; DB 9; Length 3544;
Best Local Similarity 99.8%; Pred. No. 2.1e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACGAGAGTGGTGGCTCTCTGAAA 60
DB 589 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACGAGAGTGGTGGCTCTCTGAAA 60
QY 61 AGTGAACGGAAGAAAGTGTCTAATTGATAGCCGCGCATTTTGGATACAAATACATCC 120
DB 649 AGTGAACGGAAGAAAGTGTCTAATTGATAGCCGCGCATTTTGGATACAAATACATCC 120
QY 121 CACATTTTGAAGCCATTAAATATCAACTGCTCAAGCTTTTGAAGGCAAGTTGCAACG 180
DB 709 CACATTTTGAAGCCATTAAATATCAACTGCTCAAGCTTTTGAAGGCAAGTTGCAACG 180
QY 181 GACAAAGTGTAAATACAGAGCTCATCCAGATTCAGCCGCAACATAGGTTGACATGAT 240
DB 769 GACAAAGTGTAAATACAGAGCTCATCCAGATTCAGCCGCAACATAGGTTGACATGAT 240
QY 241 TGGCTGAGAGGTTGATGATTACGATCAAGCTCCCAAGTGGTGGCTCTCTCTTCA 300
DB 829 TGGCTGAGAGGTTGATGATTACGATCAAGCTCCCAAGTGGTGGCTCTCTCTTCA 300
QY 301 GACTGTTTTCACAGTACTCTGGGTTAACTGGAGAAAGCTTCACTGTTCACTG 360
DB 889 GACTGTTTTCACAGTACTCTGGGTTAACTGGAGAAAGCTTCACTGTTCACTG 360
QY 361 CTGGCAGGTTGTTGTTGATGTTCTGCTGTTTCCCTGGGCTCTGGAGAGAAATCC 420
DB 949 CTGGCAGGTTGTTGTTGATGTTCTGCTGTTTCCCTGGGCTCTGGAGAGAAATCC 420
QY 421 ACTCTAGTCTCTACCTGCAATTTCTCAGGCTTCTTACTGTTGCCAACATTTGGGCCAAC 480
DB 1009 ACTCTAGTCTCTACCTGCAATTTCTCAGGCTTCTTACTGTTGCCAACATTTGGGCCAAC 480
QY 481 CGAATCTCTCCCAATCTTATTTTGGCTGCCAGAGATGCTCCACAGAGAGCTGAT 540
DB 1069 CGAATCTCTCCCAATCTTATTTTGGCTGCCAGAGATGCTCCACAGAGAGCTGAT 540
QY 541 CAGCAGAAATGGATGTTATGTTTAAATGCCAGCTATCTCTGCAAGAGCTGATG 1128
DB 1129 CAGCAGAAATGGATGTTATGTTTAAATGCCAGCTATCTCTGCAAGAGCTGATG 1128
QY 601 ATCCCGAGTCTCAATTTCTGCGTGGCTCTGTAATGACAGCTTTTGTAGAAAATTTTG 1188
DB 1189 ATCCCGAGTCTCAATTTCTGCGTGGCTCTGTAATGACAGCTTTTGTAGAAAATTTTG 1188
QY 661 CCGTGGTTGACAAATCAGTAGATTTCAATGAGAAAGCAAAAGCTCCCAATGATGTT 1248
DB 1249 CCGTGGTTGACAAATCAGTAGATTTCAATGAGAAAGCAAAAGCTCCCAATGATGTT 1248
QY 721 CTAGTGACATGTTTAAAGTGGATCTCCGCTCCGACCATGCTATGCTCAATCATG 1308
DB 1309 CTAGTGACATGTTTAAAGTGGATCTCCGCTCCGACCATGCTATGCTCAATCATG 1308

```

```

QY 781 AAGAGATGACATGCTTTTAAATGATGATACAGATTTGTGAAAAGAAAAGACTACT 840
DB 1369 AAGAGATGACATGCTTTTAAATGATGATACAGATTTGTGAAAAGAAAAGACTACT 840
QY 841 ATATCTCCAAACTCAATTTTCTGGGCCCACTCCGCACTATGAGAAAGATTAAGAAC 900
DB 1429 ATATCTCCAAACTCAATTTTCTGGGCCCACTCCGCACTATGAGAAAGATTAAGAAC 900
QY 901 CAGACT 906
DB 1489 CAGACT 1494

```

RESULT 8

```

US-10-377-072-25
; Sequence 25, Application US/10377072
; Publication No. US2004009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Currie, Rozy A.J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williams, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Mayoung
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 36992, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: MP103-01BOMNIM
; CURRENT APPLICATION NUMBER: US/10/377,072
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)...(2586)
US-10-377-072-25

```

```

Query Match      99.6%; Score 902.8; DB 16; Length 3544;
Best Local Similarity 99.8%; Pred. No. 2.1e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACGAGAGTGGTGGCTCTCTGAAA 60
DB 589 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACGAGAGTGGTGGCTCTCTGAAA 60

```

Qy	61	AGTGGAAAGGAAAAAGTGTCTGTATTTGATAGCCGGCCATTGTGGAATACATATCATCC	120
Db	649	AGTGGAAAGGAAAAAGTGTCTGTATTTGATAGCCGGCCATTGTGGAATACATATCATCC	708
Qy	121	CACATTTTGGAAAGCCATTAAATATCAACGTCTCAAGCTTTATGAACGGAAGGTTGCATAG	180
Db	709	CACATTTTGGAAAGCCATTAAATATCAACGTCTCAAGCTTTATGAACGGAAGGTTGCATAG	768
Qy	181	GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAAACATTAAGTTGACATTGAT	240
Db	769	GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAAACATTAAGTTGACATTGAT	828
Qy	241	TGCAGTCAGAAAGGTGTATGTTATGCATCAAAAGCTCCAAAGATGTTGCTCTCTCTTCA	300
Db	829	TGCAGTCAGAAAGGTGTATGTTATGCATCAAAAGCTCCAAAGATGTTGCTCTCTCTTCA	888
Qy	301	GACGTGTTTTCACATGTACTTCCTGGGTAACTGGAGAAAGGTTCAACTGTGTACCTG	360
Db	889	GACGTGTTTTCACATGTACTTCCTGGGTAACTGGAGAAAGGTTCAACTGTGTACCTG	948
Qy	361	CTTGCAGGTGGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGGCCCTGTGAAAGAAATACC	420
Db	949	CTTGCAGGTGGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGGCCCTGTGAAAGAAATACC	1008
Qy	421	ACTCTAGTCCCTACCTGCATTTCTCAGCCTTGCTTACCTGTGGCAACATTTGGSCAAC	480
Db	1009	ACTCTAGTCCCTACCTGCATTTCTCAGCCTTGCTTACCTGTGGCAACATTTGGSCAAC	1068
Qy	481	CGAATTTCTCCCAATCTTATCTTGGGCTGCAGAGAGATGTCCTCAAGAGAGCTGATG	540
Db	1069	CGAATTTCTCCCAATCTTATCTTGGGCTGCAGAGAGATGTCCTCAAGAGAGCTGATG	1128
Qy	541	CAGCAGATGGGATGTGTTATGTGTTAATGCGACGTAATACCTGTGCCAAAGCCTACCTT	600
Db	1129	CAGCAGATGGGATGTGTTATGTGTTAATGCGACGTAATACCTGTGCCAAAGCCTACCTT	1188
Qy	601	ATCCCGAGTCTCATTTCTCTGCGTGGCTGTGAATGAACGCTTTTGTGAGAAAATTTTG	660
Db	1189	ATCCCGAGTCTCATTTCTCTGCGTGGCTGTGAATGAACGCTTTTGTGAGAAAATTTTG	1248
Qy	661	CCGTGTTGGAACAAATCAGTAGATTTTCAITGAGAAAGCAAAAGCTCCAAATGATGTGTT	720
Db	1249	CCGTGTTGGAACAAATCAGTAGATTTTCAITGAGAAAGCAAAAGCTCCAAATGATGTGTT	1308
Qy	721	CTAAGTCACTGTTTATGCTGGGATCTCCCGCTCCGCGCACATCGCTATCGCTACATCATG	780
Db	1309	CTAAGTCACTGTTTATGCTGGGATCTCCCGCTCCGCGCACATCGCTATCGCTACATCATG	1368
Qy	781	AAGAGATGACATGTCCTTTAGATGAACCTTACAGATTTGTGAAGAAAAGACCTTACT	840
Db	1369	AAGAGATGACATGTCCTTTAGATGAACCTTACAGATTTGTGAAGAAAAGACCTTACT	1428
Qy	841	ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAGAAGATTAAAGAC	900
Db	1429	ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAGAAGATTAAAGAC	1488
Qy	901	CAGACT 906	
Db	1489	CAGACT 1494	

```

RESULT 9
US-10-425-114-26234
; Sequence 26234, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```

```

1  TITLE OR INVENTION: Plants and Uses Thereof for Plant Improvement
2
3  FILE REFERENCE: 38-21(53313)B
4
5  CURRENT APPLICATION NUMBER: US/10/425,114
6
7  CURRENT FILING DATE: 2003-04-28
8
9  NUMBER OF SEQ ID NOS: 73128
10
11 SEQ ID NO 26234
12
13 LENGTH: 3625
14
15 TYPE: DNA
16
17 ORGANISM: Homo sapiens
18
19 FEATURE:
20
21 OTHER INFORMATION: Clone ID: LIR4119-028-H6_FLI
22
23 US-10-425-114-26234

```

Query Match	99.6%	Score 902.8	DB 13	Length 3625
Best Local Similarity	99.8%	Pred. No. 2.4e-282		
Matches 904	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

QY	1	ATGGCCCATGAGATGATTTGGAACTCAATTTGTATCTGAGAGGTTGGTGGCTCTGCTGGAA	60
Db	692	ATGGCCCATGAGATGATTTGGAACTCAATTTGTATCTGAGAGGTTGGTGGCTCTGCTGGAA	751
QY	61	AGTGGAAACGAAAAAGTCTGCTAATTGATTAAGCCGGCCATTGTGGAATACATACATCC	120
Db	752	AGTGGAAACGAAAAAGTCTGCTAATTGATTAAGCCGGCCATTGTGGAATACATACATCC	811
QY	121	CACATTTTGGAAAGCATTTAATATCAACTGCTCCAAAGCTTATGAAGCCAAAGTTGCAACAG	180
Db	812	CACATTTTGGAAAGCATTTAATATCAACTGCTCCAAAGCTTATGAAGCCAAAGTTGCAACAG	871
QY	181	GACAAAGTGTTAATTACAGAGCTCATCCAGACATTCAGCGAAACATTAAGTTGACATTTGAT	240
Db	872	GACAAAGTGTTAATTACAGAGCTCATCCAGACATTCAGCGAAACATTAAGTTGACATTTGAT	931
QY	241	TGCAGTCAGAAAGTTGTAGTTTACGATCAAAGCTCCAGAAATGTGCTCTCTCTTTCA	300
Db	932	TGCAGTCAGAAAGTTGTAGTTTACGATCAAAGCTCCAGAAATGTGCTCTCTCTTTCA	991
QY	301	GACTGTTTTCTCACTGTACTTCTGGGTTAAACTGGANAAAGCTTCAACTCTGTTCACTG	360
Db	992	GACTGTTTTCTCACTGTACTTCTGGGTTAAACTGGANAAAGCTTCAACTCTGTTCACTG	1051
QY	361	CTTTCAGAGTGGGTTGGCTGAGTTCTCTCGTGTGTTCCCTGGCCCTCTGGAAGAAAAATCC	420
Db	1052	CTTTCAGAGTGGGTTGGCTGAGTTCTCTCGTGTGTTCCCTGGCCCTCTGGAAGAAAAATCC	1111
QY	421	ACTCTAGTCCCTACCTGCATTTCTCAGCTTGTCTTAACTGTGCAACATTTGGGCCAAC	480
Db	1112	ACTCTAGTCCCTACCTGCATTTCTCAGCTTGTCTTAACTGTGCAACATTTGGGCCAAC	1171
QY	481	CGAATTTCTCCCAATCTTTAATCTTGGCTGCGCAGAGAAATGTCCTCAACAGAGACTGATG	540
Db	1172	CGAATTTCTCCCAATCTTTAATCTTGGCTGCGCAGAGAAATGTCCTCAACAGAGACTGATG	1231
QY	541	CAGCAGATGAGGATTTGTTATGTGTTAAATGCGCAGCTATACCTGTGCCAAAGCCTGACTTT	600
Db	1232	CAGCAGATGAGGATTTGTTATGTGTTAAATGCGCAGCTATACCTGTGCCAAAGCCTGACTTT	1291
QY	601	ATCCCCCAGTCTCAATTTCTGCTGCTGCTGCTGTGAATGACAGCTTTTGTGAGAAAATTTTG	660
Db	1292	ATCCCCCAGTCTCAATTTCTGCTGCTGCTGCTGTGAATGACAGCTTTTGTGAGAAAATTTTG	1351
QY	661	CCGTTGGTTGGAACAAATCAGTAAAGATTTCAATTGAGAAACAAAGGCTCCAAATGAGATGTT	720
Db	1352	CCGTTGGTTGGAACAAATCAGTAAAGATTTCAATTGAGAAACAAAGGCTCCAAATGAGATGTT	1411
QY	721	CTAGTGCACGTGTTAGCTGGGAATCTCCCGCTCCGCAACAATCGCTATCGCTCAATCATG	780
Db	1412	CTAGTGCACGTGTTAGCTGGGAATCTCCCGCTCCGCAACAATCGCTATCGCTCAATCATG	1471
QY	781	AAGAGATGCAATGCTTTTATGATGAAGCTTACAGATTTTGTGAAAAGAAAAGACTACT	840
Db	1472	AAGAGATGCAATGCTTTTATGATGAAGCTTACAGATTTTGTGAAAAGAAAAGACTACT	1531

QY 841 ATATCTCCAACTTCAATTTCTGGGCGCACTCTCGACATATGAGAAAGATTAAAGAC 900
DB 1532 ATATCTCCAACTTCAATTTCTGGGCGCACTCTCGACATATGAGAAAGATTAAAGAC 1591
QY 901 CAGACT 906
DB 1592 CAGACT 1597

RESULT 10
US-10-343-357-17
; Sequence 17, Application US/10343357
; Publication No. US20040058341A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.Tom
; APPLICANT: ELIOTT, Vicki S.; RAMKUMAR, Jayalakmi
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
; APPLICANT: LEE, Ernestine A.; HAPALIA, April J.A.
; APPLICANT: LU, Dying Aina M.; TRIBOLEY, Catherine M.
; APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
; APPLICANT: YUB, Henry; WARREN, Bridget A.
; APPLICANT: NGUYEN, Damien B.; CHAWLA, Narinder K.
; APPLICANT: KEARNEY, Liam
; TITLE OF INVENTION: PROTEIN PHOSPHATASES
; FILE REFERENCE: PI-0173 PCT
; CURRENT APPLICATION NUMBER: US/10/343,357
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: PCT/US01/23716
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,679
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/223,272
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/224,309
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 60/226,728
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/229,254
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: US 60/231,366
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 3766
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CBI
US-10-343-357-17

Query Match 99.6%; Score 902.8; DB 13; Length 3766;
Best Local Similarity 99.8%; Pred. No. 2,1e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGATGATTTGAATCAATTTGTTACTGAGAGTTGGTGGCTGCTGGA 60
DB 538 ATGGCCCATGATGATTTGAATCAATTTGTTACTGAGAGTTGGTGGCTGCTGGA 597
QY 61 AGTGAAGGAAAAGCTGCTTAATGATGAGCGGCAATTTGTGATATACATATCATCC 120
DB 598 AGTGAAGGAAAAGCTGCTTAATGATGAGCGGCAATTTGTGATATACATATCATCC 657
QY 121 CACATTTTGAAGCATTAATATCACTGCTCAAGCTTATAGAGCAAGGTTGCAAG 180
DB 658 CACATTTTGAAGCATTAATATCACTGCTCAAGCTTATAGAGCAAGGTTGCAAG 717
QY 181 GACAAAGTTAATTCAGAGCTCATCCGCAATTCAGCGAAACATTAAGTTGACATTTGAT 240
DB 718 GACAAAGTTAATTCAGAGCTCATCCGCAATTCAGCGAAACATTAAGTTGACATTTGAT 777

QY 241 TGCAGTCAGAGGTTGTAGTTTACATCAAGCTCCAGAGATGTGGCTCTCTCTTCA 300
DB 778 TGCAGTCAGAGGTTGTAGTTTACATCAAGCTCCAGAGATGTGGCTCTCTCTTCA 837
QY 301 GACTGTTTCTCACTGTATCTTCTGGGTAACTGAGAAAGACTTCACTCTGTTACCTG 360
DB 838 GACTGTTTCTCACTGTATCTTCTGGGTAACTGAGAAAGACTTCACTCTGTTACCTG 897
QY 361 CTGGAGGTGGGTTTGTGAGATCTCTGCTGTTGTTTCTCTGCTCTGTTGAGAAATCC 420
DB 898 CTGGAGGTGGGTTTGTGAGATCTCTGCTGTTGTTTCTCTGCTCTGTTGAGAAATCC 957
QY 421 ACTCTAGTCCCTACCTGATTTCTGACCTTGTCTACTGTTGCAACATTTGGCCAAAC 480
DB 958 ACTCTAGTCCCTACCTGATTTCTGACCTTGTCTACTGTTGCAACATTTGGCCAAAC 1017
QY 481 CGAATCTTCCCAATCTTTATCTTGTGCTGCGAGAGATGCTCTTAACAAGAGCTGATA 540
DB 1018 CGAATCTTCCCAATCTTTATCTTGTGCTGCGAGAGATGCTCTTAACAAGAGCTGATA 1077
QY 541 CAGAGAAATGGATGTTATGTTAAATGCGACCTATACCTGTCGAAGCTGACTTT 600
DB 1078 CAGAGAAATGGATGTTATGTTAAATGCGACCTATACCTGTCGAAGCTGACTTT 1137
QY 601 ATCCCGAGTCTATTTCTGCGGTGTCGCTGTAATGACAGCTTTTGTGAGAAATTTTG 660
DB 1138 ATCCCGAGTCTATTTCTGCGGTGTCGCTGTAATGACAGCTTTTGTGAGAAATTTTG 1197
QY 661 CGGTGTTGGAACAATCAGTATTTCAATGAGAAAGCAAAAGCTTCAATGATGTTG 720
DB 1198 CGGTGTTGGAACAATCAGTATTTCAATGAGAAAGCAAAAGCTTCAATGATGTTG 1257
QY 721 CTATGCACTGTTTGTAGCTGGATCTCCGCGTCCGCAATCTGTTATGCTTACATCATG 780
DB 1258 CTATGCACTGTTTGTAGCTGGATCTCCGCGTCCGCAATCTGTTATGCTTACATCATG 1317
QY 781 AAGAGATGAGACATGCTTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
DB 1318 AAGAGATGAGACATGCTTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1377
QY 841 ATATCTCCAACTTCAATTTCTGGGCGCACTCTCGACATATGAGAAAGATTAAAGAC 900
DB 1378 ATATCTCCAACTTCAATTTCTGGGCGCACTCTCGACATATGAGAAAGATTAAAGAC 1437
QY 901 CAGACT 906
DB 1438 CAGACT 1443

RESULT 11
US-10-648-593-115
; Sequence 115, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 115
; LENGTH: 4790
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-648-593-115

Query Match 99.6%; Score 902.8; DB 17; Length 4790;

Best Local Similarity 99.8%; Pred. No. 2,5e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATGATGGAAGTCAATTTGTTACTGAGAGTGTGGCTGTCTGGAA 60
DB 184 ATGGCCCATGATGATGGAAGTCAATTTGTTACTGAGAGTGTGGCTGTCTGGAA 243
QY 61 AGTGAACGGAAGAGTGTGATGATGAGAGTGTGGCTGTCTGGAAATCAATCACTCC 120
DB 244 AGTGAACGGAAGAGTGTGATGATGAGAGTGTGGCTGTCTGGAAATCAATCACTCC 303
QY 121 CACATTTTGAAGCATTAATATCAATGCTCCAGCTTAATGAAGCAAGTTGCAACAG 180
DB 304 CACATTTTGAAGCATTAATATCAATGCTCCAGCTTAATGAAGCAAGTTGCAACAG 363
QY 181 GACAAAGTGTAAATTAACAGCTCATCAGCATTCAGCGAAATTAAGTTGACATTGAT 240
DB 364 GACAAAGTGTAAATTAACAGCTCATCAGCATTCAGCGAAATTAAGTTGACATTGAT 423
QY 241 TGCAGTCAGAAAGTGTGATGATGAGAGTGTGGCTGTCTGGAAATCAATCACTCC 300
DB 424 TGCAGTCAGAAAGTGTGATGATGAGAGTGTGGCTGTCTGGAAATCAATCACTCC 483
QY 301 GACTGTTTCTCACTGATCTTGGGTAACAGTGAAGAGCTTCACTGTTCACCTG 360
DB 484 GACTGTTTCTCACTGATCTTGGGTAACAGTGAAGAGCTTCACTGTTCACCTG 543
QY 361 CTTCAGAGTGGGTTTGTGATGATCTCTGTTGTTTCCCTGGCCTGTGTGAAGAAATCC 420
DB 544 CTTCAGAGTGGGTTTGTGATGATCTCTGTTGTTTCCCTGGCCTGTGTGAAGAAATCC 603
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTACCTGTGCAACATTTGGCCCAAC 480
DB 604 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTACCTGTGCAACATTTGGCCCAAC 663
QY 481 CGAATTTCTTCCCAATCTTTATCTTGGCTGCGAGAGATGCTCTCAACAAGAGCTGATA 540
DB 664 CGAATTTCTTCCCAATCTTTATCTTGGCTGCGAGAGATGCTCTCAACAAGAGCTGATA 723
QY 541 GACGAGATGGAGTGTGATGATGATTAATGCAAGCTTAATCTGTCCAAAGCTGACTTT 600
DB 724 GACGAGATGGAGTGTGATGATGATTAATGCAAGCTTAATCTGTCCAAAGCTGACTTT 783
QY 601 ATCCCGAGTCAATTTCTCAGCCTGCTGTAATGCAAGCTTGTGTGAAGAAATTTTG 660
DB 784 ATCCCGAGTCAATTTCTCAGCCTGCTGTAATGCAAGCTTGTGTGAAGAAATTTTG 843
QY 661 CCGTGTGTGACAATCATGATGATTTCAATGAGAAAGCAAAAGCTTCAATGATGTGTT 720
DB 844 CCGTGTGTGACAATCATGATGATTTCAATGAGAAAGCAAAAGCTTCAATGATGTGTT 903
QY 721 CTAGTGACTGTTTAACTGGGATCTCCGCTCGCCCAACATGCTATGCTTCAATCATG 780
DB 904 CTAGTGACTGTTTAACTGGGATCTCCGCTCGCCCAACATGCTATGCTTCAATCATG 963
QY 781 AAGAGATGACATGCTTAAATGAGCTTAAGATTTGTAAGAAAGAAAGAAAGCTTACT 840
DB 964 AAGAGATGACATGCTTAAATGAGCTTAAGATTTGTAAGAAAGAAAGAAAGCTTACT 1023
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGTGACTATGAGAAAGATTAAGAAC 900
DB 1024 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGTGACTATGAGAAAGATTAAGAAC 1083
QY 901 CAGACT 906
DB 1084 CAGACT 1089

RESULT 12
US-10-094-749-673
; Sequence 673, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:

APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKI, ICHIRO
APPLICANT: SEKI, NAOTIKO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKYU
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIRO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 673
LENGTH: 2102
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-673

Query Match 99.5%; Score 901.2; DB 16; Length 2102;
Best Local Similarity 99.7%; Pred. No. 4.9e-282;
Matches 903; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATGATGGAAGTCAATTTGTTACTGAGAGTGTGGCTGTCTGGAA 60
DB 56 ATGGCCCATGAGATGATGGAAGTCAATTTGTTACTGAGAGTGTGGCTGTCTGGAA 115
QY 61 AGTGAACGGAAGAGTGTGATGATGAGAGTGTGGCTGTCTGGAAATCAATCACTCC 120
DB 116 AGTGAACGGAAGAGTGTGATGATGAGAGTGTGGCTGTCTGGAAATCAATCACTCC 175
QY 121 CACATTTTGAAGCATTAATATCAATGCTCCAGCTTAATGAAGCAAGTTGCAACAG 180
DB 176 CACATTTTGAAGCATTAATATCAATGCTCCAGCTTAATGAAGCAAGTTGCAACAG 235
QY 181 GACAAAGTGTAAATTAACAGCTCATCAGCATTCAGCGAAATTAAGTTGACATTGAT 240
DB 236 GACAAAGTGTAAATTAACAGCTCATCAGCATTCAGCGAAATTAAGTTGACATTGAT 295
QY 241 TGCAGTCAGAAAGTGTGATGATGATTAATGCAAGCTCCCAAGATGTGCTCTCTTCA 300
DB 296 TGCAGTCAGAAAGTGTGATGATGATTAATGCAAGCTCCCAAGATGTGCTCTCTTCA 355
QY 301 GACTGTTTCTCACTGATCTTGGGTAACAGTGAAGAGCTTCACTGTTCACCTG 360
DB 356 GACTGTTTCTCACTGATCTTGGGTAACAGTGAAGAGCTTCACTGTTCACCTG 415
QY 361 CTTCAGAGTGGGTTTGTGATGATCTCTGTTGTTTCCCTGGCCTGTGTGAAGAAATCC 420
DB 416 CTTCAGAGTGGGTTTGTGATGATCTCTGTTGTTTCCCTGGCCTGTGTGAAGAAATCC 475
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTACCTGTGCAACATTTGGCCCAAC 480
DB 476 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTACCTGTGCAACATTTGGCCCAAC 535
QY 481 CGAATTTCTTCCCAATCTTTATCTTGGCTGCGAGAGATGCTCTCAACAAGAGCTGATA 540
DB 536 CTAAATCTTCCCAATCTTTATCTTGGCTGCGAGAGATGCTCTCAACAAGAGCTGATA 595

```
QY 541 CAGCAGATGGATTGCTTATGTGTTAAATGACGCTATACCTGTCCAAAGCCGACTTT 600
Db 596 CAGCAGATGGATTGCTTATGTGTTAAATGACGCTATACCTGTCCAAAGCCGACTTT 655
QY 601 ATCCCGAGTCTCATTTCTCGGTGTGCTGTGTAATGACGCTTTGTGGAATAATTTTG 660
Db 656 ATCCCGAGTCTCATTTCTCGGTGTGCTGTGTAATGACGCTTTGTGGAATAATTTTG 715
QY 661 CCGTGTGGGACAAATCGATGATTTTCAATGAGAAAGCAAAAGCTTCAATGATGATTT 720
Db 716 CCGTGTGGGACAAATCGATGATTTTCAATGAGAAAGCAAAAGCTTCAATGATGATTT 775
QY 721 CTAGTGCATGTTTAACTGTGGATCTCCGCTCCGCAACCAATGCTATGCTTACATCATG 780
Db 776 CTAGTGCATGTTTAACTGTGGATCTCCGCTCCGCAACCAATGCTATGCTTACATCATG 835
QY 781 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAAGACCTTACT 840
Db 836 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAAGACCTTACT 895
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGTAATGAGAAAGATTTAAGAAC 900
Db 896 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGTAATGAGAAAGATTTAAGAAC 955
QY 901 CAGACT 906
Db 956 CAGACT 961
```

RESULT 13

```
US-10-296-115-520
; Sequence 520, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: 2002-11-18
; PRIOR APPLICATION NUMBER: US/10/296,115
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 520
; LENGTH: 2966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2966)
; OTHER INFORMATION: n = a,t,c or g
US-10-296-115-520
```

Query Match

Best Local Similarity 99.3%; Score 899.6; DB 13; Length 2966;

Matches 902; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 ATGGCCCATGAGATGATGGAACCAATTGTTTACTGAGAGGTGCTGCTGTGAA 60
Db 23 ATGGCCCATGAGATGATGGAACCAATTGTTTACTGAGAGGTGCTGCTGTGAA 82
QY 61 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCCGCAATTTGTGGAATPAAATCATCC 120
Db 83 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCCGCAATTTGTGGAATPAAATCATCC 142
QY 121 CACATTTTGAAGCATTATATCACTGCTCAAGCTTATGACGGAAGTTTGAACAG 180
Db 143 CACATTTTGAAGCATTATATCACTGCTCAAGCTTATGACGGAAGTTTGAACAG 202
QY 181 GACAAAGTTTAAATGACGCTCATCGACATTTGAGGAAACATTAAGTTGACATTTGAT 240
```

```
Db 203 GACAAAGTTTAAATGACGCTCATCGACATTTGAGGAAACATTAAGTTGACATTTGAT 262
QY 241 TGAAGTCAGAAAGTTGTGATTTTACGATCAAGCTCCCAAGATGTTGCTCTCTTCA 300
Db 263 TGAAGTCAGAAAGTTGTGATTTTACGATCAAGCTCCCAAGATGTTGCTCTCTTCA 322
QY 301 GACTGTTTCTCAGTGTCTTCTGGGTAAATGGAAGAGCTTCAACTGTTCACCTG 360
Db 323 GACTGTTTCTCAGTGTCTTCTGGGTAAATGGAAGAGCTTCAACTGTTCACCTG 382
QY 361 CTGGAAGGTGTTTGTCTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAGGAATTC 420
Db 383 CTGGAAGGTGTTTGTCTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAGGAATTC 442
QY 421 ACTTGAATCCCTACCTGATTTCTCAGCCCTGTTACCTGTTGCCAATTTGGGCAACC 480
Db 443 ACTTGAATCCCTACCTGATTTCTCAGCCCTGTTACCTGTTGCCAATTTGGGCAACC 502
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCCAGAGATGCTTCAACAGAGCTGATA 540
Db 503 CGAATTTCTCCCAATCTTTATCTTGGCTGCCAGAGATGCTTCAACAGAGCTGATA 562
QY 541 CAGCAAGTGAATGTTGTTATGTTAAATGCACTTACTCTGCCAAGCTTACT 600
Db 563 CAGCAAGTGAATGTTGTTATGTTAAATGCACTTACTCTGCCAAGCTTACT 622
QY 601 ATCCCGAGTCTCATTTTCTGCTGTGCTGTGAAATGACAGCTTTGTGAAAAATTTTG 660
Db 623 ATCCCGAGTCTCATTTTCTGCTGTGCTGTGAAATGACAGCTTTGTGAAAAATTTTG 682
QY 661 CCGTGTGGAACAAATCGATGATTTTCAATGAGAAAGCAAAAGCTTCAATGATGTT 720
Db 683 CCGTGTGGAACAAATCGATGATTTTCAATGAGAAAGCAAAAGCTTCAATGATGTT 742
QY 721 CTAGTGCATGTTTAACTGTGGATCTCCGCTCCGCAACCAATGCTTACATCATG 780
Db 743 CTAGTGCATGTTTAACTGTGGATCTCCGCTCCGCAACCAATGCTTACATCATG 802
QY 781 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAAGACCTTACT 840
Db 803 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAAGACCTTACT 862
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGTAATGAGAAAGATTTAAGAAC 900
Db 863 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGTAATGAGAAAGATTTAAGAAC 922
QY 901 CAGACT 906
Db 923 CAGACT 928
```

RESULT 14

US-10-072-012-255

; Sequence 255, Application US/10072012

; Publication No. US2004003493A1

; GENERAL INFORMATION:

```
; APPLICANT: Tchernov, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Eshe
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Guev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
```



```

APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsebrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 255
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
US-10-072-012-255

```

Query Match	90.3%	Score 817.8	DB 13	Length 2200
Best Local Similarity	95.1%	Pred. No. 7.5e-25.5		
Matches	904	Conservative	0	Mismatches 2
				Indels 45
				Gaps 4

QY	1	ATGGCCCATGAGATGATTGGAATCTCAATTTGTATTCTGAGAGGTTGATGGCTCTGCTGGA	60
Db	61	ATGGCCCATGAGATGATTGGAATCTCAATTTGTATTCTGAGAGGTTGATGGCTCTGCTGGA	120
QY	61	AGTGAACGGAAAAAGTCTGCTTAATTGATAGCCGGCCATTGTGGAATCAATTCATCC	120
Db	121	AGTGAACGGAAAAAGTCTGCTTAATTGATAGCCGGCCATTGTGGAATCAATTCATCC	180
QY	121	CACATTTTGGAAAGCCATTAAATCAACGCTCCAGGCTTATGAAACGGAAGTTGGAACAG	180
Db	181	CACATTTTGGAAAGCCATTAAATCAACGCTCCAGGCTTATGAAACGGAAGTTGGAACAG	240
QY	181	GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGGAAACATA-----AG	228
Db	241	GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGGAAACATAAGGTAAAGCTCAG	300
QY	229	GTTGACATTGATTGCAATCAGAGGTTGATGTTACGATCAAAAGCTCCCAAGATGTTGCC	288
Db	301	GTTGACATTGATTGCAATCAGAGGTTGATGTTACGATCAAAAGCTCCCAAGATGTTGCC	360
QY	289	TCTGCTCTTGAGACTGTTTTCTCACTGTAATTTCTGGGTTAACTGGAGAAAGGCTTCAAC	348
Db	361	TCTGCTCTTGAGACTGTTTTCTCACTGTAATTTCTGGGTTAACTGGAGAAAGGCTTCAAC	420
QY	349	TCTGTTCACTGCTTGC-----AGGTGGGTTTGCTGAATTCCTCGATGTTTC	396
Db	421	TCTGTTCACTGCTTGCAGAGTTTAATCTTAAGTGGGTTTGCTGAATTCCTCGATGTTTC	480
QY	397	CTGGGCTCTGTGAAGAAAAATCACTCTAATGTCCTACTGCAATTTCTAGGCTTGTCTTA	456

Db	481	CTGGCCCTCTGTGAAGAAATTCACCTCACTGACCTTCACTGACATTTCTCAGCCTTGTCTTA	540
Oy	457	CTGTGTGGCAACATTTGGGCCAAACCCGGAATTTCTTCCCAATCTTTATCTTGGCTGGCAGCGA	516
Db	541	CTGTGTGGCAACATTTGGGCCAAACCCGGAATTTCTTCCCAATCTTTATCTTGGCTGGCAGCGA	600
Oy	517	GATGTCTTCCAACAAGAGACTGATACAGACAGATGGGATGGTATGTGTATTAATGCAAGC	576
Db	601	GATGTCTTCCAACAAGAGACTGATACAGACAGATGGGATGGTATGTGTATTAATGCAAGC	660
Oy	577	TATATCTGTCCAAAGCTTGACTTTATCCCCGAGTCTCATTTCTCTGCGTGTGCTGTGAAT	636
Db	661	AATATCTGTCCAAAGCTTGACTTTATCCCCGAGTCTCATTTCTCTGCGTGTGCTGTGAAT	720
Oy	637	GACAGCTTTTGTGAGAAATTTTGGCGTGTGAGCAATCATGATGATTTCAAT	690
Db	721	GACAGCTTTTGTGAGAAATTTTGGCGTGTGAGCAATCATGATGATTTCAATGCTTAG	780
Oy	691	-----GAGAAAGCAAAAGCCTTCCAATGGAATGTGTTCTATGTGACATGTTTAGCT	738
Db	781	TTGACATTATACAGAGAAAGCAAAAGCCTTCCAATGGAATGTGTTCTATGTGACATGTTTAGCT	840
Oy	739	GGGATCTTCCGCTCCGACCACCATGCTATGCGCTCATCATGAAGAGATGACATGTCT	798
Db	841	GGGATCTTCCGCTCCGACCACCATGCTATGCGCTCATCATGAAGAGATGACATGTCT	900
Oy	799	TTAAGATGAAGCTTAC--AGATTGTGAAAGAAAAAGACCTATCATATCTTCCAACTTTC	855
Db	901	TTAAGATGAAGCTTACAGAGATTTGTGTAAAGAAAAAGACCTATCATATCTTCCAACTTTC	960
Oy	856	AATTTTTCGGGCAAACTCTGTGACATATAGAAAGAAATTAAGAACACAGACT	906
Db	961	AATTTTTCGGGCAAACTCTGTGACATATAGAAAGAAATTAAGAACACAGACT	1011

```

RESULT 15
US-09-964-277-20
; Sequence 20, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucite, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-964-277-20

Query Match          62.3%; Score 564.8; DB 9; Length 3332;
Best Local Similarity 81.7%; Pred. No. 2.6e-172;
Matches 740; Conservative 0; Mismatches 2; Indels 164; Gaps 1.

QY      1 ATGCCCATGATGATGATTGGAATCTCAAAATTGTACTGAGAGGTTGGTGCTCTGCTGGA 60
DB      562 ATGCCCATGATGAGATTTGGAACTCAAAATTGTACTGAGAGGTTGGTGCTCTGCTGGA 621

QY      61 AGTGAACGGAAAAAGTGTCTTAATTGATAGCGGCGCATTTGTGGAATACATACATCC 120
DB      622 AGTGAACGGAAAAAGTGTCTTAATTGATAGCGGCGCATTTGTGGAATACATACATCC 681

QY      121 CACATTTTGAAGGCATTAATATCACTGCTCCAGGTTATGAAGCGAAGTTGCAACAG 180
DB      682 CACATTTTGAAGGCATTAATATCACTGCTCCAGGTTATGAAGCGAAGTTGCAACAG 741

QY      181 GACAAATGTTAATTACAGAGCTCATTCAGATTCACGAAACATAAGGTTGAATTCAT 240
DB      742 GACAAATGTTAATTACAGAGCTCATTCAGATTCACGAAACATAAGGTTGAATTCAT 801

```

```
QY 241 TGCAGTCAGAGGTTGTACTTTACGATAAAGCTCCCAAGATGTGCTCTCTCTCA 300
Db |||||
Db 802 TGCAGTCAGAGGTTGTACTTTACGATCAAGCTCCCAAGATGTGCTCTCTCTCA 861
QY 301 GACTGTTTTCTCACTGTACTCTGTGGTAACTGAGAAAGCTTCAACTCTGTCACTG 360
Db |||||
Db 862 GACTGTTTTCTCACTGTACTCTGTGGTAACTGAGAAAGCTTCAACTCTGTCACTG 921
QY 361 CTTCAGGTTGGGTTTGTGAGTCTCTGTGTTTCCCTGCTCTGTGAAGAAATCC 420
Db |||||
QY 922 CTTCG----- 926
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCTTGTCTTACCTGTTCGCAACATTGGGCCAAC 480
Db ----- 926
QY 927 ----- 926
Db 481 CGAATCTTCCCAATCTTTATCTTGCTGCCAGCAGATGCTTCACAGAGAGCTGATA 540
QY 927 ----- -AGGAGCTGATG 937
Db 541 CAGCAGATGGGATGGTATGTATGTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 600
QY 938 CAGCAGATGGGATGGTATGTATGTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 660
Db |||||
QY 601 ATCCCGAGTCTCATTTCTGCTGTGCTGTGTAATGACAGCTTTGTGAGAAATTTTG 660
Db |||||
QY 998 ATCCCGAGTCTCATTTCTGCTGTGCTGTGTAATGACAGCTTTGTGAGAAATTTTG 1057
QY 661 CCGTGGTTGCAAAATCAGTGAATTTCAATTGAGAAAGCAAAAGCTTCGAATGATGTT 720
Db |||||
QY 1058 CCGTGGTTGCAAAATCAGTGAATTTCAATTGAGAAAGCAAAAGCTTCGAATGATGTT 1117
QY 721 CTAGTCACTGTTTACGTGGGATCTCCCGCTCCGCCACATGCTATGCTTACATCAG 780
Db |||||
QY 1118 CTAGTCACTGTTTACGTGGGATCTCCCGCTCCGCCACATGCTTACATCAG 1177
QY 781 AAGAGATGACATGCTCTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 840
Db |||||
QY 1178 AAGAGATGACATGCTCTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 1237
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGACTATGAGAGAGATTAGAAC 900
Db |||||
QY 1238 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGACTATGAGAGAGATTAGAAC 1297
QY 901 CAGACT 906
Db |||||
QY 1298 CAGACT 1303
```

Search completed: June 22, 2004, 03:25:21
Job time : 543.914 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 17:08:00 ; Search time 74.9535 Seconds
(without alignments)
6707.969 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443
Perfect score: 906
Sequence: 1 atggcccatgagatgatg.....agaagataagaaccagact 906

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
 - 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
 - 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
 - 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
 - 5: /cgn2_6/ptodata/2/ina/PCOMB.COMB.seq:*
 - 6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	902.8	99.6	1998	4 US-09-816-494-3	Sequence 3, Appl1
2	902.8	99.6	3544	4 US-09-816-494-1	Sequence 1, Appl1
3	365.8	40.4	2377	4 US-09-920-668-3	Sequence 3, Appl1
4	223	24.6	279	4 US-09-016-434-91	Sequence 91, Appl1
5	124.6	13.8	1830	4 US-09-557-921-1	Sequence 1, Appl1
6	106.8	11.8	2303	4 US-09-922-146-3	Sequence 1, Appl1
7	96.2	10.6	2109	4 US-09-016-434-1135	Sequence 1135, App
8	96.2	10.6	2109	4 US-09-023-655-946	Sequence 946, App
9	94.6	10.4	1208	4 US-09-023-655-347	Sequence 347, App
10	91.6	10.1	1619	4 US-09-702-705-801	Sequence 801, App
11	91.6	10.1	1619	4 US-09-736-457-801	Sequence 801, App
12	91.6	10.1	1619	4 US-09-614-1248-801	Sequence 801, App
13	91.6	10.1	1619	4 US-09-614-1248-801	Sequence 801, App
14	91.6	10.1	1619	4 US-09-589-184-801	Sequence 801, App
15	91.6	10.1	4637	4 US-09-702-705-804	Sequence 804, App
16	91.6	10.1	4637	4 US-09-736-457-804	Sequence 804, App
17	91.6	10.1	4637	4 US-09-614-1248-804	Sequence 804, App
18	91.6	10.1	4637	4 US-09-614-1248-804	Sequence 804, App
19	91.6	10.1	4637	4 US-09-589-184-804	Sequence 804, App
20	90	9.9	1238	2 US-08-530-290-11	Sequence 11, Appl
21	90	9.9	1238	4 US-09-702-705-803	Sequence 803, App
22	90	9.9	1238	4 US-09-736-457-803	Sequence 803, App
23	90	9.9	1238	4 US-09-614-1248-803	Sequence 803, App
24	90	9.9	1238	4 US-09-614-1248-803	Sequence 803, App
25	90	9.9	1238	4 US-09-589-184-803	Sequence 803, App
26	90	9.9	2064	4 US-09-702-705-825	Sequence 825, App
27	90	9.9	2064	4 US-09-736-457-825	Sequence 825, App

28	90	9.9	2064	4 US-09-614-1248-825	Sequence 825, App
29	90	9.9	2064	4 US-09-671-325-825	Sequence 825, App
30	90	9.9	2064	4 US-09-589-184-825	Sequence 825, App
31	90	9.9	2109	4 US-09-702-705-826	Sequence 826, App
32	90	9.9	2109	4 US-09-736-457-826	Sequence 826, App
33	90	9.9	2109	4 US-09-614-1248-826	Sequence 826, App
34	90	9.9	2109	4 US-09-671-325-826	Sequence 826, App
35	90	9.9	2109	4 US-09-589-184-826	Sequence 826, App
36	90	9.9	2240	4 US-09-016-434-1100	Sequence 1100, App
37	87	9.6	2000	4 US-09-016-434-1291	Sequence 1291, App
38	86.2	9.5	240	4 US-09-016-434-776	Sequence 776, App
39	80.4	8.9	944	4 US-09-371-671B-10	Sequence 10, Appl1
40	80.4	8.9	1987	2 US-08-990-379-1	Sequence 1, Appl1
41	80.4	8.9	1993	2 US-08-990-379-2	Sequence 2, Appl1
42	79	8.7	539	4 US-09-389-681-311	Sequence 311, App
43	79	8.7	539	4 US-09-620-405B-311	Sequence 311, App
44	79	8.7	539	4 US-09-339-338-311	Sequence 311, App
45	79	8.7	539	4 US-09-433-826B-311	Sequence 311, App

ALIGNMENTS

RESULT 1					
US-09-816-494-3					
Sequence 3, Application US/09816494					
Patent No. 6664089					
GENERAL INFORMATION:					
APPLICANT: Meyers, Rachel A.					
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY					
TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR					
FILE REFERENCE: 10448-030002 US/09/816,494					
CURRENT FILING DATE: 2001-03-23					
PRIOR APPLICATION NUMBER: US 60/191,858					
PRIOR FILING DATE: 2000-03-24					
NUMBER OF SEQ ID NOS: 10					
SOFTWARE: FastSeq for Windows Version 4.0					
SEQ ID NO 3					
LENGTH: 1998					
TYPE: DNA					
ORGANISM: Homo sapiens					
US-09-816-494-3					
Query Match					
Best Local Similarity 99.8%; Score 902.8; DB 4; Length 1998;					
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
QY	1	ATGGCCATGAGATGATGGAATGTTACTGAGAGTGGTGGCTGCTGGAA	60		
DB	1	ATGGCCATGAGATGATGGAATGTTACTGAGAGTGGTGGCTGCTGGAA	60		
QY	61	AGTGAACGGAAGATGCTGCTAATGATAGCCGCAATTTGTGAAATCAATATCC	120		
DB	61	AGTGAACGGAAGATGCTGCTAATGATAGCCGCAATTTGTGAAATCAATATCC	120		
QY	121	CACATTTGGAACCATTAATCAACTGCTCAAGCTTATGAAGCAAGTTGCAACAG	180		
DB	121	CACATTTGGAACCATTAATCAACTGCTCAAGCTTATGAAGCAAGTTGCAACAG	180		
QY	181	GACAAAGTGTATTAACAGAGCTCATCGACATTCAGCGAAATATAGGTTGACATGAT	240		
DB	181	GACAAAGTGTATTAACAGAGCTCATCGACATTCAGCGAAATATAGGTTGACATGAT	240		
QY	241	TGCAATGAGAGGTTGATTTACATCAAGCTCCCAAGATGTCCTCTCTTCA	300		
DB	241	TGCAATGAGAGGTTGATTTACATCAAGCTCCCAAGATGTCCTCTCTTCA	300		
QY	301	GATGTTTCTCCTGCTACTCTTGGGTAATGAGAAAGCTTCAACTCTGTTCACTG	360		
DB	301	GATGTTTCTCCTGCTACTCTTGGGTAATGAGAAAGCTTCAACTCTGTTCACTG	360		
QY	361	CTTGCAAGTGGTGGTCTGATGTTCTCTGTTGTTCTCTGCTCTGTAAGAAATCC	420		

Db 361 CTTGACAGTGGTTTGCTGAGTTCTCTGTTTCCCTGGCTCTGAGAAATCC 420
Qy 421 ACTCTAGTCCCTTCTGCAATTTCTGAGCTTCTTACCGTTGCAACATTTGGCCAAAC 480
Db 421 ACTCTAGTCCCTTCTGCAATTTCTGAGCTTCTTACCGTTGCAACATTTGGCCAAAC 480
Qy 481 CGAATTTCTTCCCAATCTTATCTTGGCTGCGCAAGAGATGCTCTCAACAGAGCTGATA 540
Db 481 CGAATTTCTTCCCAATCTTATCTTGGCTGCGCAAGAGATGCTCTCAACAGAGCTGATA 540
Qy 541 CAGCAAGATGGGATTTGTTATGTTTAAATGCGACTATACCTGTCCAAAGCTGACTTT 600
Db 541 CAGCAAGATGGGATTTGTTATGTTTAAATGCGACTATACCTGTCCAAAGCTGACTTT 600
Qy 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 CCGTGGTTGCACAATGATGATGATTTGATGAGAAAGCAAGCCCTCAATGAGATGTT 720
Db 661 CCGTGGTTGCACAATGATGATGATTTGATGAGAAAGCAAGCCCTCAATGAGATGTT 720
Qy 721 CTATGACCTGTTTAACTGAGTCTCCGCTCCGCAACATGCTATGCTATGCTATCATG 780
Db 721 CTATGACCTGTTTAACTGAGTCTCCGCTCCGCAACATGCTATGCTATGCTATCATG 780
Qy 781 AAGAGATGACATGCTTATGATGAGCTTACAGATTTGAGAAAGAAAGCACTACT 840
Db 781 AAGAGATGACATGCTTATGATGAGCTTACAGATTTGAGAAAGAAAGCACTACT 840
Qy 841 ATATCTCCAACTTCAATTTCTGCGGCACTCCGCACTATGAGAAAGATTAAGAAC 900
Db 841 ATATCTCCAACTTCAATTTCTGCGGCACTCCGCACTATGAGAAAGATTAAGAAC 900
Qy 901 CAGACT 906
Db 901 CAGACT 906

RESULT 2
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: PHOSPHATASE MOLECULES AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2583)
US-09-816-494-1

Query Match 99.6%; Score 902.8; DB 4; Length 3544;
Best Local Similarity 99.8%; Pred. No. 3, 3e-309;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATGGAAGTCAATTTTACTGAGAGTGGTGGTCTGCTGGA 60
Db 589 ATGGCCCATGAGATGATGGAAGTCAATTTTACTGAGAGTGGTGGTGGTCTGCTGGA 60
Qy 61 AATGGAACGAAAGTCTGCTTAATTGATAGCCGCACTTTTGGAATACATATCCTC 120

Db 649 AGTGAACGAAAAAGTGTGCTTAATTGATACCGGCACTTTTGGAATACATATCCTC 708
Qy 121 CACATTTTGGAAACCATTAATATCAAGCTCTCAAGCTTATGAAAGCGAGTTGCAACAG 180
Db 709 CACATTTTGGAAACCATTAATATCAAGCTCTCAAGCTTATGAAAGCGAGTTGCAACAG 180
Qy 181 GACAAAGTGTAAATTAACAGCTCTATCCAGATTCAGCAAACTAATAGTTGACATTTGAT 240
Db 769 GACAAAGTGTAAATTAACAGCTCTATCCAGATTCAGCAAACTAATAGTTGACATTTGAT 240
Qy 241 TCGAGTCAGAGTGTGATTTAAGATCAAAAGCTCCCAAGATGCTCTCTCTTCA 300
Db 829 TCGAGTCAGAGTGTGATTTAAGATCAAAAGCTCCCAAGATGCTCTCTCTTCA 300
Qy 301 GACTGTTTCTACGTATCTTGGGTAACCTGAGAAAGCTTCAACTCTGTCACCTG 360
Db 889 GACTGTTTCTACGTATCTTGGGTAACCTGAGAAAGCTTCAACTCTGTCACCTG 360
Qy 361 CTTGACAGTGGTGTGCTGAGTCTCTGTTGTTCCCTGGCTCTGAGAAAGAAATCC 420
Db 949 CTTGACAGTGGTGTGCTGAGTCTCTGTTGTTCCCTGGCTCTGAGAAAGAAATCC 420
Qy 421 ACTCTAGTCCCTTCTGCAATTTCTGAGCTTCTTACCTGTTGCAACATTTGGCCAAAC 480
Db 1009 ACTCTAGTCCCTTCTGCAATTTCTGAGCTTCTTACCTGTTGCAACATTTGGCCAAAC 480
Qy 481 CGAATTTCTTCCCAATCTTATCTTGGCTGCGCAAGATGCTCTCTCTCTCTTCA 540
Db 1069 CGAATTTCTTCCCAATCTTATCTTGGCTGCGCAAGATGCTCTCTCTCTCTTCA 540
Qy 541 CAGCAAGATGGGATTTGTTATGTTTAAATGCGACTATACCTGTCCAAAGCTGACTTT 600
Db 1129 CAGCAAGATGGGATTTGTTATGTTTAAATGCGACTATACCTGTCCAAAGCTGACTTT 600
Qy 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 1189 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 CCGTGGTTGCACAATGATGATGATTTGATGAGAAAGCAAGCCCTCAATGAGATGTT 720
Db 1249 CCGTGGTTGCACAATGATGATGATTTGATGAGAAAGCAAGCCCTCAATGAGATGTT 720
Qy 721 CTATGACCTGTTTAACTGAGTCTCCGCTCCGCAACATGCTATGCTATGCTATCATG 780
Db 1309 CTATGACCTGTTTAACTGAGTCTCCGCTCCGCAACATGCTATGCTATGCTATCATG 780
Qy 781 AAGAGATGACATGCTTATGATGAGCTTACAGATTTGAGAAAGAAAGCACTACT 840
Db 1369 AAGAGATGACATGCTTATGATGAGCTTACAGATTTGAGAAAGAAAGCACTACT 840
Qy 841 ATATCTCCAACTTCAATTTCTGCGGCACTCCGCACTATGAGAAAGATTAAGAAC 900
Db 1429 ATATCTCCAACTTCAATTTCTGCGGCACTCCGCACTATGAGAAAGATTAAGAAC 900
Qy 901 CAGACT 906
Db 1489 CAGACT 1494

RESULT 3
US-09-920-668-3
; Sequence 3, Application US/09920668
; Patent No. 6482644
; GENERAL INFORMATION:
; APPLICANT: Brett P. Cowart
; APPLICANT: Lex M. Cowart
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 8 EXPRESSION
; FILE REFERENCE: RUS-0246
; CURRENT APPLICATION NUMBER: US/09/920,668
; PRIOR FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3

LENGTH: 2377
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (135) ... (2012)
 US-09-920-668-3

Query Match 40.4%; Score 365.8; DB 4; Length 2377;
 Best Local Similarity 64.5%; Pred. No. 8.8e-119;
 Matches 563; Conservative 0; Mismatches 307; Indels 3; Gaps 1;

28 ATTGTTACTGAGAGTTGGTGGCTCTGCTGGAAGTGAACGGAAAAAGTGGCTGCAATT 87
 165 ATGGATGCCAAGAGCTGGCCAGCTCTGCGGGGCGGGGCGGGGCTGGTATC 224
 88 GATAGCGGCGCATTTTGGAAATACATATACATCCCATTTTGGAGCCATTATATGAC 147
 225 GACAGCGGCTCCTTCGTGAGATACACAGCTGGCATGTGCTCAGCTCCGTCAATCTGC 284
 148 TGCTCCAGCTTATGAAGCGAAGTTGCAACAGCAAAAGTTTATTAATACAGCTATC 207
 285 TGCTCCAGCTTGGAAAGCGGCTGACAGGCAAGGTGACCATTTGGAGCTTATC 344
 208 CAGCATTCAGGAAACATTAAGTTGACATTGATTGACATGAGAGTTGATTATGAT 267
 345 CAGCGGCTGACGACGCGAGGTGAGGCTACGAGCCACAGACGTGTGCTATGAC 404
 268 CAAAGCTCCCAAGATTTGCTCTCTCTTCAAGATGTTTCTCACTGTAATCTTGGGT 327
 405 CAGAGCAGCGGGAGCGCAGGCTGTGCGCAGACAGCTTCTCTCATCTGCTGAC 464
 328 AAATGGAAGAGCTTCAACTCTGTTCACTGCTTCAAGTGGGTTTCTGAGTTCT 387
 465 AAGCTGACCGCTGCTTCAAGGCTGCTCTCACTGCGGCTTCTGACCTTCTCC 524
 388 CGTTGTTCCCTGGCTCTGGAAGAAATTCACCT---CTAGCCCTTCACTGATTTCT 444
 525 TCCTGCTTCCCGGCTCTGAGAGGAGCAGCTGCTCCCTGCTACCAAGAGCTTCTCC 584
 445 CAGCTTGTACTGCTTGGCAATTTGGGCAACCCGAAATTTCCCAATCTTTATCT 504
 585 CAGCGCTGCTGCTGAGGCTGAGGCTGACCGGCAATCTGCTCACTCTATCTG 644
 505 GCGTCCAGGAGATGCTCTCAACAGAGCTGATACAGCAATGGGATTTGTTATG 564
 645 GCGTCCGAGAGGAGCTCTAAACAGATCTGATGACGCAAAATGGAATAGCTATGTC 704
 565 TTAAATGCGAGCTATACCTGCTCAAAAGCTGACTTTATCCCGAGTCTCATTTCCGCT 624
 705 CTCACCGCAGCACTCTGCCCCAAGCTGACTTCACTGCGAGAGCCGCTTCATGCGG 764
 625 GTGCTGTGATGACAGCTTTTGTGAAAAATTTGCGGTGTGCAAAATCAGTATG 684
 765 GTCCCATCAACGACACTCTGTGAAAACTGCTGCGGCTGCGGCAAGTCCATGAG 824
 685 TTCAATTGAAGAAAGCAAGCTTCAATGATGTGTTTCTAGTCACTGTTTACTGAGATC 744
 825 TTCAATCATTAAGCAAGCTCTCAAGCTGCCAAGTCTATGCTCACTGTCTGCGGATC 884
 745 TCCGCTCGGCAACATGCTATGCTTACATCATGAGAGAGATGACATGTCTTATGAT 804
 885 TCCGCTCTGCAACATGCTGCTATCACTATGAGACCAATGGGCAATGTCTCCGAC 944
 805 GAAGCTTACAGATTTGTGAAGAAAAAGACTTATCTATCTCCAAATTTCTG 864
 945 GAGCCTACAGGTTCTGGAAGGACAGGCGCCGTCATCTGCGCAATCTCAATCTCTG 1004
 865 GCGCACTCTGGAATGAGAGAAAGATTAG 897
 1005 GCGCAGCTGTGAGTACAGCGCAGCTGAG 1037

RESULT 4

US-09-016-434-91

Sequence 91, Application US/09016434

Patent No. 6500938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

TITLE OF INVENTION: PATHWAY GENE EXPRESSION

NUMBER OF SEQUENCES: 1490

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

FILING DATE: US/09/016,434

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEO ID NO: 91:

SEQUENCE CHARACTERISTICS:

LENGTH: 279 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: LUNGFET03

CLONE: 1234795

US-09-016-434-91

Query Match 24.6%; Score 223; DB 4; Length 279;

Best Local Similarity 100.0%; Pred. No. 8.5e-69;

Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCATGAGATGATGGAATGGAATCTCAATTTTACTGAGAGTTGCTGCTGGAA 60

DB 25 ATGGCCATGAGATGATGGAATCTCAATTTTACTGAGAGTTGCTGCTGGAA 84

QY 61 AGTGAACGAAAGAGTGTGCTTAATTTGATAGCGGCACTTTTGTGAATCAATACATCC 120

DB 85 AGTGAACGAAAGAGTGTGCTTAATTTGATAGCGGCACTTTTGTGAATCAATACATCC 144

QY 121 CACATTTGGAAGCCATTATATCACTGCTCAAGCTTATGAAGCAAGTTGCAACAG 180

DB 145 CACATTTGGAAGCCATTATATCACTGCTCAAGCTTATGAAGCAAGTTGCAACAG 204

QY 181 GACAAAGTTTAATTAAGAAGCTCATCAGCATTCAGCGAAC 223

DB 205 GACAAAGTTTAATTAAGAAGCTCATCAGCATTCAGCGAAC 247

RESULT 5

US-09-557-921-1

Sequence 1, Application US/09557921

Patent No. 6551810

GENERAL INFORMATION:
APPLICANT: Luche, Ralf M.
TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.416
CURRENT APPLICATION NUMBER: US/09/557,921
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1830
TYPE: DNA
ORGANISM: Homo sapiens
US-09-557-921-1

Query Match
Best Local Similarity 13.8%; Score 124.6; DB 4; Length 1830;
Matches 251; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
QY 461 TTGCCAATGTTGGGCAACCCGAAATTTCTTCCCAATTTTATCTGGCTGCCAGGAGATG 520
DB 1036 TCGAAGACGCTGAGCTCAACCCCAATCTTGCCCTTCTGCTTGGCAATGAGGAGATG 1095
QY 521 TCTCAACAGAGAGCTGATACAGCAGATGAGATGTTATGTATGTTAATGCCAGTATA 580
DB 1096 CTGAGGACCTGACACCATGACCGGCTGACATGGCTAGCTATCAACCTCACACTC 1155
QY 581 CTTGTCGAAGCTGACTT---ATCCCGAGTCTCATTTCTCGTGGTGGCTGTGATG 637
DB 1156 ATCTTCCCTCTACCTATGAGAAAGCTGTTCATCAACGAGCTGCGACCACTG 1215
QY 638 ACAGCTTTTGTGAAAAATTTTGGCGTGTGACCAATGATGATTTTATTGAGAAAG 697
DB 1216 ACAGCAACAGAGAACCTGCGGCTGACTTTGAAAGGCTTTTGAAGTATTGAGAAAG 1275
QY 698 CAAAAGCTTCAATGATGATGTTCTAGTGCATCTTTAGCTGGAGATCTCCGCTCCGCCA 757
DB 1276 CTCACCAAGTGGGAGAGGCGTTCTCATCTCCAGCTGGGAGTGTCCGCTCCGCCA 1335
QY 758 CCATGCTATGCGCTCAATCATATGAGAGAGATGATGTTTATGATGAGATTTACAGAT 817
DB 1336 CCATGCTATGCTTACTTATGAGAGACACCTCGAGTACATGATGATGCTTATTAAT 1395
QY 818 TTGTGAAAGAAAAAGACTACTATATCTCCCAACTTCAATTTTGTGGCCCACTCTGG 877
DB 1396 TTGTCAAGGCAACGACGCAATATCTCCCAACTTCAATTTGAGGAGATTTGCTAG 1455
QY 878 ACTATGAGAGAGATTAAGAAC 900
DB 1456 AGTTGAGAGAGACTTAACACAC 1478

RESULT 6
US-09-922-146-3
Sequence 3, Application US/09922146
Patent No. 6566133
GENERAL INFORMATION:
APPLICANT: Lex M. Cowser
APPLICANT: Brett P. Monia
TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION
CURRENT APPLICATION NUMBER: US/09/922,146
CURRENT FILING DATE: 2001-08-01
NUMBER OF SEQ ID NOS: 48
SEQ ID NO 3
LENGTH: 2303
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (114)...(1268)
US-09-922-146-3

Query Match
Best Local Similarity 11.8%; Score 106.8; DB 4; Length 2303;
Matches 229; Conservative 0; Mismatches 182; Indels 3; Gaps 1;
QY 475 CCAACCGAATTTCTTCCCAATCTTATCTTGGCTGCCAGGAGATGTTCTCAACAGAG 534
DB 723 CTGTCCAGATCTTCCCAACTTATCTTGGGAGTGGCCGAGATTCGCAATTGGAG 782
QY 535 CTGATACAGAGAAATGGATTTGTTATGTAAATGCACTTACCTGTCCAA---G 591
DB 783 AGCTGCCAACTGGGAGATCCGCTACATCTCATATGCAACCCCAACTCCCAACTTC 842
QY 592 CCGACTTATTCGCCAGTCTATTTCTCGTGGTGGCTGTGAATGACAGCTTTGTAG 651
DB 843 TTGGAAGAAATGAGATTTCTTCACTAAGCAATCCCAATCTCCAGCCTGAGCGCG 902
QY 652 AAAATTTTGGCGTGTGACCAATCATGATGATTTGATGAGAAAGCAAAAGCTTCCAA 711
DB 903 AACCTGTGCGGCTTTTCCGAGGCGCATTTGATGATGAGGCTTGTCCAGAAC 962
QY 712 GGATGTGTTTATGCACTGTTTACGTTGAGTCTCCGCTCCGACCATGCTATGCGC 771
DB 963 TCGGGGTGCTGTGCACTGCTTGGCGGGGTACGCTTCTGTACCGTCACTGTGCGC 1022
QY 772 TACATCATGAGAGAGATGAGCATGCTTTAATGAGACTTACAGATTTGTGAAGAAAA 831
DB 1023 TACCTATGAGAGAGCTTCACTCTCTTCAACGATGCTATGATGCTGTGCAAGAGAG 1082
QY 832 AGACCTATATCTTCAAACTTCAATTTCTGGGCGCAACTCGGACTATGAG 885
DB 1083 AAGCTTAAGATCTCCCAACTTCACTTATGAGGAGGAGTGTGAGCTTGTAG 1136

RESULT 7
US-09-016-434-1135
Sequence 1135, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1135:
SEQUENCE CHARACTERISTICS:
LENGTH: 2109 base pairs

```

; TYPE: nucleic acid
; STRANDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1418933
US-09-016-434-1135

```

Query Match	10.6%	Score 96.2	DB 4	Length 2109
Best Local Similarity	53.8%	Pred. No. 3.8e-23		
Matches 222, Conservative	0	Mismatches 188	Indels 3	Gaps 1

Qy	484	ATTCTTCCCAATCTTTATCTTGTCGTCGACGAGATGTCCTCAACAAGAGCTGTATACAG	543
Db	979	ATCTTGCCCTTCCTCTACTTGCGCTGTGCAAAAGCTCACCMACTTGAGACGTGTTGAG	1033
Qy	544	CAGAAATGGATGGTTATGTATGTTAAATGCCAGCTATACCTGTCCAAAGC--CTGACTTT	600
Db	1039	GAATTCGGCATCAAGTATCATCTTGAACCTCACCCCAATTGGCCGAATCTCTTTAGAAC	1098
Qy	601	ATCCCCGAGTCTCATTTTCTGCGTGTGCTGTGAATGACAGTTTGTGAGAAATTTTG	660
Db	1099	GCAGAGAGGTTTAAATACAGCAAAATCCCACTGGATACATGAGGCCAAACCTGTCC	1158
Qy	661	CCGTGGTTGACAAATCAGTATGATTTCTATTGAGAAAGCAAAAGCTCCAAATGATGTGT	720
Db	1159	CAGTTTTTCCCTGAGGCCATTTCTTCAATGATGAAGCCGGGGCAAGAACTGTGTGTCTC	1218
Qy	721	CTAGTGACCTGTTTAGCTGGGATCTCCGCTCCGCGCACATCGCTATCGCTCAATCATG	780
Db	1219	TTGGTACATTTGCTGGCTGGCATTTAGCCGCTCAGTCACTGTGATCTGTGGCTTACTATG	1278
Qy	781	AAGAGATGACATGTCTTTAGATBAAGCTTACAGATTTGTCAAGAAAAAGAAAGCTTACT	840
Db	1279	CAGAACTCAATCTGTGCATGAAACATGCTATGACATTTGTCMAAATGAAAAAATCCAAC	1338
Qy	841	ATTATCCCAAACTTCAATTTTTCTGGGGCAACCTCGGACCTATGAGAAAGAT	893
Db	1339	ATATTCCTTAATTCACTTCAATGGGTGACGTCTGAGACTTTCGAGAGAGCGCT	1391

RESULT 8
 US-09-023-655-946
 : Sequence 946, Application US/09023655
 : Patent No. 6607879
 :
 GENERAL INFORMATION:
 :
 APPLICANT: Cocks, Benjamin G.
 APPLICANT: Susan G. Stuart
 APPLICANT: Jeffrey J. Sellhammer
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENOME
 TITLE OF INVENTION: EXPRESSION
 NUMBER OF SEQUENCES: 1508
 :
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 STREET: 3174 PORTER DRIVE
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94304
 :
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: word Perfect 6.1 for Windows/MS-DOS 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/023,655
 FILING DATE: HERewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:

```

1 NAME: Zeller, Karen J.
2 REGISTRATION NUMBER: 37,071
3 REFERENCE/DOCKET NUMBER: PA-0001 US
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: (650) 855-0555
6 TELEFAX: (650) 845-4166
7 INFORMATION FOR SEQ. ID NO: 946:
8 SEQUENCE CHARACTERISTICS:
9 LENGTH: 2109 base pairs
10 TYPE: nucleic acid
11 STRANDEDNESS: single
12 TOPOLOGY: linear
13 IMMEDIATE SOURCE:
14 LIBRARY: GENBANK
15 CLONE: G1418933
16
17 US-09-023-655-946

```

Query Match	10.6%	Score 96.2	DB 4	Length 2109
Best Local Similarity	53.8%	Pred. No. 3.8e-23		
Matches 222	Conservative 0	Mismatches 188	Indels 3	Gaps 1

Qy	484	ATTCTTCCCAATCTTTATCTTGGCTGCGCAGGAGATGCTTCAACAGAGACTGTATACAG	543
Db	979	ATCTTGCCCTTCCCTCACTTGGGCTGTGGCCAAAGCTCCACCAACTTGGAGCTGTGTGAG	1033
Qy	544	CAGATGGGATGGTTATGTGTAAATGCGACGTATACCTGTCCAAAGC---CTGACTTT	600
Db	1039	GAATTCGGCATCAAGTACATCTTGAGAGCTACCCCAATTTGCGGAATCTCTTGTGAGAAC	1098
Qy	601	ATCCCGAGTCATTTCTCGCGTGTGCTGTGAATGACACTTTTGTGAGAAAATTTTG	660
Db	1099	GCAGAGAGTTTAATACAGCAAAATCCCATCTCGATCACTGAGCAAAACCTGTCC	1155
Qy	661	CCGTGTGTGAGCAAAATCAGTAGATTTCTATTGAGAAAGCAAAAGCCTCAATGATGTGTT	720
Db	1159	CAGTTTTTCCCTGAGGCCATTTCTTTCAATGATGAAGCCGGGCGAAGACTGTGTGTCC	1218
Qy	721	CTAGTGCATGTTTAACTGGGATCTCCGCTCCGCCACATCGTATCGCTACATCATG	780
Db	1219	TTGTATACATTGCTTTGGCTGGCATTAAGCCGCTCAATGACTGTGACTGTGGCTTACTTATG	1278
Qy	781	AAGAGATGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAAAGAAAAGACTACT	840
Db	1279	CAGAACTCATCTGTGATGAAGAGATGCTATGACATTTGTCAAAATGAAAAATTCACAC	1338
Qy	841	ATATCTCCAAACTCAATTTTCTGGGCGCAACTCTGACATATGGAAGAAAT	893
Db	1339	ATATCTCCCTTAATCTCAATCTCATAGGATGCTGTGTGACTGTGAGATCTGAGAGGACCT	1391

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/023,655
 FILING DATE: HEREMITH

CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:

CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.
 REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 347:
 SEQUENCE CHARACTERISTICS:

LENGTH: 1208 base pairs
 TYPE: nucleic acid

STRANDEDNESS: single
 TOPOLOGY: linear

IMMEDIATE SOURCE:
 LIBRARY: THKNOT03

CLONE: 1444245
 US-09-023-655-347

Query Match 10.4%; Score 94.6; DB 4; Length 1208;
 Best Local Similarity 64.3%; Pred. No. 9.3e-23;
 Matches 142; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 680 TAGATTTTCAATGAGAAAGCAAAAGCTTCAATGATGTGTTAGAGCACTGTTAGCTG 739
 DB 4 TTGAGTTTATTTGAG 739
 QY 740 GGAATCTCCGCTCCGAG 739
 DB 64 GGGGTGCTCCGCTCCGAG 739
 QY 800 TAGATGAGCTTACAGATTTTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
 DB 124 TGACTGATGCTTAAATTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
 QY 860 TTCTGGGCAACTCTGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 DB 184 TCATGGGAGTGTGCTGAG 224

RESULT 10

US-09-702-705-801
 Sequence 801, Application US/09702705
 Patent No. 6504010

GENERAL INFORMATION:
 APPLICANT: Wang, Tonglong

APPLICANT: Bangur, Chaitanya S.
 APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary
 APPLICANT: Vedvick, Tom

APPLICANT: Carter, Darick
 APPLICANT: Retter, Marc

APPLICANT: Mannion, Jane
 APPLICANT: Fan, Liqun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.478C14
 CURRENT APPLICATION NUMBER: US/09/702,705

CURRENT FILING DATE: 2000-10-30
 NUMBER OF SEQ ID NOS: 1833

SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 801

LENGTH: 1619
 TYPE: DNA

ORGANISM: Homo sapien
 US-09-702-705-801

Query Match 10.1%; Score 91.6; DB 4; Length 1619;
 Best Local Similarity 53.2%; Pred. No. 1.3e-21;
 Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

QY 484 ATTCTTCCCAATCTTATCTTGGCTGCGAGAGATGCTTCAACAAGAGAGCTGATACG 543
 DB 320 ATCTTCTCCCTCTTCTTACCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
 QY 544 CAGATGGAATTTGTTATGTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
 DB 380 GCGCTGAGATCAAGGCTGTTGATGTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 436
 QY 604 CCGAGTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
 DB 437 GAGACTTACATGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 496
 QY 664 TGTGTCGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 723
 DB 497 TGTGTCGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 556
 QY 724 GTGACTGTTTACCTGAGATCTTCCGCTCCGACCATGCTATGCTTACATGATGATGATG 783
 DB 557 GTGACTGTCAG 616
 QY 784 AGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 843
 DB 617 AAGAAAG 893
 QY 844 TCTCAACTTCAATTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 DB 677 TCCCAACTTCAATTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726

RESULT 11

US-09-736-457-801
 Sequence 801, Application US/09736457
 Patent No. 6509448

GENERAL INFORMATION:
 APPLICANT: Wang, Tonglong

APPLICANT: Bangur, Chaitanya S.
 APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary
 APPLICANT: Vedvick, Tom

APPLICANT: Carter, Darick
 APPLICANT: Retter, Marc

APPLICANT: Mannion, Jane
 APPLICANT: Fan, Liqun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.478C15
 CURRENT APPLICATION NUMBER: US/09/736,457

CURRENT FILING DATE: 2000-12-13
 NUMBER OF SEQ ID NOS: 1864

SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 801

LENGTH: 1619
 TYPE: DNA

ORGANISM: Homo sapien
 US-09-736-457-801

Query Match 10.1%; Score 91.6; DB 4; Length 1619;
 Best Local Similarity 53.2%; Pred. No. 1.3e-21;
 Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

QY 484 ATTCTTCCCAATCTTATCTTGGCTGCGAGAGATGCTTCAACAAGAGAGCTGATACG 543
 DB 320 ATCTTCTCCCTCTTCTTACCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
 QY 544 CAGATGGAATTTGTTATGTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
 DB 380 GCGCTGAGATCAAGGCTGTTGATGTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 436

Qy 604 CCGAGCTCATTTCTCGCGTGGCTGTAATGACAGCTTTTGTGAAAATTTTGGCG 663
Db 437 GGACATATACATGACAGATGATCCAGTGGAGATACCAAGGCCGACATGACTCC 496
Qy 664 TGGTTGACAATACATGATATTTTCATTTAGAAAAGCAAGCTTCATATGATGTTCTA 723
Db 497 TGGTTATGAGAGCATAGATGATCATGATCGCTGAAAGATGCTGGGGCGGTGCTG 556
Qy 724 GTGCACTGTTTACGTGGGATCTCCGCTCCGCCACCATGCTATGCTTACATCATGAAG 783
Db 557 GTGCACTGCCAGGGGGCATCTCGCGCTCGCCACCATCTGCGCTTACCTGATGATG 616
Qy 784 AGATGACATGCTTATGATGAACTTACAGATTTGTGAAAAGAAAAGACCTACTATA 843
Db 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTGAGTTCTTTAAGACGCCGCCGACATTAATC 676
Qy 844 TCTCCAACTTCAATTTTCTGGGCCAACTCTGACTGATGAGAAAGAT 893
Db 677 TCGCCCACTTCAAGCTTCAATGGGGCAGCTGCTGAGTTGAGTCCAGGT 726

RESULT 12
US-09-614-124B-801
; Sequence 801, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-801

Query Match 10.1%; Score 91.6; DB 4; Length 1619;
Best Local Similarity 53.2%; Pred. No. 1.3e-21;
Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
Qy 484 ATCTTCCCAATCTTATCTTGTGCTGCACGAGATGCTCTCAACAAGAGCTGATACAG 543
Db 320 ATCTTCCCTCTCTCACTCGGAGTGCCTACATGCTGCGGAGAGACATGCTGAC 379
Qy 544 CAGAATGGATGTTGATGTTAATGCACTATACCTGTCCAAAGCTGACTTATC 603
Db 380 GCCCTGGGCAATCAAGGCTCTGTGAATGCTCTCGGACATGCCAAAC--ACTTTGAA 436
Qy 604 CCGAGCTCATTTCTCGGTGCTGCTGTAATGACAGCTTTTGTGAAAATTTTGGCG 663
Db 437 GGACATATACATGACAGATGATCCAGTGGAGATACCAAGGCCGACATGACTCC 496
Qy 664 TGGTTGACAATACATGATATTTTCATTTAGAAAAGCAAGCTTCATATGATGTTCTA 723
Db 497 TGGTTATGAGAGCATAGATGATCATGATCGCTGAAAGATGCTGGGGCGGTGCTG 556
Qy 724 GTGCACTGTTTACGTGGGATCTCCGCTCCGCCACCATGCTATGCTTACATCATGAAG 783
Db 557 GTGCACTGCCAGGGGGCATCTCGCGCTCGCCACCATCTGCGCTTACCTGATGATG 616
Qy 784 AGATGACATGCTTATGATGAACTTACAGATTTGTGAAAAGAAAAGACCTACTATA 843

Db 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTCAATTTGTTAAGACAGGCCGACATTAATC 676
Qy 844 TCTCCAACTTCAATTTTCTGGGCCCACTCTGACTGATGAGAAAGAT 893
Db 677 TCGCCCACTTCAAGCTTCAATGGGGCAGCTGCTGAGTTGAGTCCAGGT 726

RESULT 13
US-09-671-325-801
; Sequence 801, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-801

Query Match 10.1%; Score 91.6; DB 4; Length 1619;
Best Local Similarity 53.2%; Pred. No. 1.3e-21;
Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
Qy 484 ATCTTCCCAATCTTATCTTGTGCTGCACGAGATGCTCTCAACAAGAGCTGATACAG 543
Db 320 ATCTTCCCTCTCTCACTCGGAGTGCCTACATGCTGCGGAGAGACATGCTGAC 379
Qy 544 CAGAATGGATGTTGATGTTAATGCACTATACCTGTCCAAAGCTGACTTATC 603
Db 380 GCCCTGGGCAATCAAGGCTCTGTGAATGCTCTCGGACATGCCAAAC--ACTTTGAA 436
Qy 604 CCGAGCTCATTTCTCGGTGCTGCTGTAATGACAGCTTTTGTGAAAATTTTGGCG 663
Db 437 GGACATATACATGACAGATGATCCAGTGGAGATACCAAGGCCGACATGACTCC 496
Qy 664 TGGTTGACAATACATGATATTTTCATTTAGAAAAGCAAGCTTCATATGATGTTCTA 723
Db 497 TGGTTATGAGAGCATAGATGATCATGATCGCTGAAAGATGCTGGGGCGGTGCTG 556
Qy 724 GTGCACTGTTTACGTGGGATCTCCGCTCCGCCACCATGCTATGCTTACATCATGAAG 783
Db 557 GTGCACTGCCAGGGGGCATCTCGCGCTCGCCACCATCTGCGCTTACCTGATGATG 616
Qy 784 AGATGACATGCTTATGATGAACTTACAGATTTGTGAAAAGAAAAGACCTACTATA 843
Db 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTCAATTTGTTAAGACAGGCCGACATTAATC 676
Qy 844 TCTCCAACTTCAATTTTCTGGGCCCACTCTGACTGATGAGAAAGAT 893
Db 677 TCGCCCACTTCAAGCTTCAATGGGGCAGCTGCTGAGTTGAGTCCAGGT 726

RESULT 14
US-09-589-184-801
; Sequence 801, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:

```

; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-589-184-801
```

```

Query Match      10.1%; Score 91.6; DB 4; Length 1619;
Best Local Similarity 53.2%; Pred. No. 1.3e-21;
Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
```

```

QY 484 ATCTTCCCAATCTTTATCTTGCTGCCAGAGATGTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTCTACTCTGCGAGATGCTTACCAAGCTGCCGAGAGACATGCTGAC 379
QY 544 CAGATGGAGATGGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTTATC 603
DB 380 GCCCTGGACATCAGGCTGTGTAATGTCTCTCGAGCTGCCAACC---ACTTTGAA 436
QY 604 CCCGAGTCTATTTCTGCGTGGCTGAGTGAAGACAGCTTTTGTGAAAAATTTTCCG 663
DB 437 GGAACATATCAGTACAGTGCATCCAGTGAAGATACCAAGGCCACATCAGCTCC 496
QY 664 TGGTTGACAAATCAGTATGATTTTCAATGAGAAAGCAAAAGCTCCAAATGATGTGCTA 723
DB 497 TGGTTCAATGAGACCAATAGATCATGATGCCGTGAAGAGACTGCGTGGGCGCTG 556
QY 724 GTGCACGTGTTTACCTGAGATCTCCGCTCCGCCACCATGCTATGCTTACATCATGAG 783
DB 557 GTGCACGTGCGAGCGGAGCATCTGCGGTGCGCCACCATCTGCTGCTTACCTGATGATG 616
QY 784 AGATGACATGCTCTTTAGATGAGCTTACAGATTTGTGAAAAAGAAAGCACTACTATA 843
DB 617 AAGAAACGGGTGAGCTGAGAGAGGCTTCGAGTTCTGTTAAGCAGCGCCGACATTAATC 676
QY 844 TCTCCAACTTCATTTTCTGGGCCAATCTCTGACATGAGAAAGAT 893
DB 677 TCGCCCACTTCACTTCAATGAGGAGGAGCTGCTGACGTTGAGTCCAGGT 726
```

```

RESULT 15
US-09-702-705-804
; Sequence 804, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
```

```

; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 804
; LENGTH: 4637
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-804
```

```

Query Match      10.1%; Score 91.6; DB 4; Length 4637;
Best Local Similarity 53.2%; Pred. No. 2.9e-21;
Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
```

```

QY 484 ATCTTCCCAATCTTTATCTTGCTGCCAGAGATGTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTCTACTCTGCGAGATGCTTACCAAGCTGCCGAGAGACATGCTGAC 379
QY 544 CAGATGGAGATGGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTTATC 603
DB 380 GCCCTGGACATCAGGCTGTGTAATGTCTCTCGAGCTGCCAACC---ACTTTGAA 436
QY 604 CCCGAGTCTATTTCTGCGTGGCTGAGTGAAGACAGCTTTTGTGAAAAATTTTCCG 663
DB 437 GGAACATATCAGTACAGTGCATCCAGTGAAGATACCAAGGCCACATCAGCTCC 496
QY 664 TGGTTGACAAATCAGTATGATTTTCAATGAGAAAGCAAAAGCTCCAAATGATGTGCTA 723
DB 497 TGGTTCAATGAGACCAATAGATCATGATGCCGTGAAGAGACTGCGTGGGCGCTG 556
QY 724 GTGCACGTGTTTACCTGAGATCTCCGCTCCGCCACCATGCTATGCTTACATCATGAG 783
DB 557 GTGCACGTGCGAGCGGAGCATCTGCGGTGCGCCACCATCTGCTGCTTACCTGATGATG 616
QY 784 AGATGACATGCTCTTTAGATGAGCTTACAGATTTGTGAAAAAGAAAGCACTACTATA 843
DB 617 AAGAAACGGGTGAGCTGAGAGAGGCTTCGAGTTCTGTTAAGCAGCGCCGACATTAATC 676
QY 844 TCTCCAACTTCATTTTCTGGGCCAATCTCTGACATGAGAAAGAT 893
DB 677 TCGCCCACTTCACTTCAATGAGGAGGAGCTGCTGACGTTGAGTCCAGGT 726
```

```

Search completed: June 22, 2004, 02:56:09
Job time : 75.9535 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 21:02:42 ; Search time 7805.33 Seconds
(without alignments)
3692.750 Million cell updates/sec

Title: US-10-029-345A-109
Perfect score: 3418
Sequence: 1 MAHEMTGTQVTERLVALLE.....LGKVGSSSTRSGSMETLEIVS 665

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xjh
-Q/cgna2.1/USPTO.spool/US10029345/runat.21062004.164146.29030/app.query.fasta.1.1294
-DB=GenBml -QPMF=fastap -SUFFIX=rgs -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10029345.OCGN.1.1.6617 @runat.21062004.164146.29030 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6.
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenBml:.*
1: gb_ba:.*
2: gb_htg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_stc:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*
15: em_ba:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_mu:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pat:.*
24: em_ph:.*
25: em_pl:.*
26: em_ro:.*
27: em_stc:.*
28: em_un:.*

29: em_vl:.*
30: em_htg_hum:.*
31: em_htg_in:.*
32: em_htg_other:.*
33: em_htg_mus:.*
34: em_htg_pln:.*
35: em_htg_rod:.*
36: em_htg_mam:.*
37: em_htg_vit:.*
38: em_sy:.*
39: em_htg_hum:.*
40: em_htg_mus:.*
41: em_htg_other:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3418	100.0	5450	6	AX482439 Sequence
2	3418	100.0	5450	6	AX482478 Sequence
3	3406	99.6	1998	6	AX260342 Sequence
4	3406	99.6	2732	6	AX180875 Sequence
5	3406	99.6	3059	6	AX278461 Sequence
6	3406	99.6	3104	6	AX405700 Sequence
7	3406	99.6	3496	6	AX441210 Sequence
8	3406	99.6	3521	9	AB052156 Sequence
9	3406	99.6	3544	6	AX260340 Sequence
10	3406	99.6	3566	6	AF506796 Sequence
11	3406	99.6	3766	6	AX374994 Sequence
12	3406	99.6	4790	6	BD171157 Novel gen
13	3406	99.6	4790	6	BD183422 Novel gen
14	3406	99.6	4790	6	AB051487 Homo sapi
15	3399	99.4	2102	6	AX713989 Sequence
16	3399	99.4	2102	6	AK055973 Sequence
17	3388	99.1	5111	6	AX482372 Sequence
18	3379.5	98.9	2071	6	AX921917 Sequence
19	3379.5	98.9	3284	9	BC042101 Homo sapi
20	3373	98.7	2118	6	AX099933 Sequence
21	3358.5	98.3	2200	6	AX921915 Sequence
22	3163	92.5	2807	9	BC031643 Homo sapi
23	3096.5	90.6	4943	10	BC059232 Mus muscu
24	3096.5	90.6	4975	10	BC057321 Mus muscu
25	3079.5	90.1	4874	10	AB052157 Mus muscu
26	3057.5	89.5	2756	6	AX482444 Sequence
27	3048.5	89.2	1935	9	AY038927 Homo sapi
28	3048.5	89.2	3332	6	AX441229 Sequence
29	3017.5	88.3	4992	10	AF345951 Mus muscu
30	2676	78.3	4827	10	AF345952 Mus muscu
31	2016	59.0	17206	9	AC007619 Homo sapi
32	2016	59.0	18834	2	AC131617 Homo sapi
33	1733.5	50.7	20147	2	AC126692 Mus muscu
34	1733.5	50.7	23659	2	AC118035 Mus muscu
35	1721	50.4	24250	2	AC133722 Rattus no
36	1721	50.4	24605	2	AC097818 Rattus no
37	1721	50.4	25078	2	AC128093 Rattus no
38	1643.5	48.1	1916	6	AX835305 Sequence
39	1643.5	48.1	1916	6	AK098310 Homo sapi
40	1505.5	44.0	4026	10	AF345953 Mus muscu
41	1422.5	41.6	3231	5	BC063941 Homo sapi
42	1327	38.8	2264	9	BC045110 Homo sapi
43	1326	38.8	2377	6	AR255944 Sequence
44	1326	38.8	2377	9	HSU27193 Human prote
45	1304	38.2	2600	10	BC052705 Mus muscu

RESULT 1

ALIGNMENTS

AX482439
 LOCUS AX482439 5450 bp DNA linear PAT 16-AUG-2002
 DEFINITION Sequence 108 from Patent WO2057460.
 AX482439
 VERSION AX482439.1 GI:22316984
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Toderud C.G., Bol D., Finger, J., Lee, L., Nelson, T., Schieven, G., Mcatee, S., Banas, D., Bassolino, D., Feder, J., Krystek, S., Polynucleotides encoding human phosphatase.
 JOURNAL Polynucleotides encoding human phosphatase.
 SOUTHB BRISTOL MYERS CO (US)
 FEATURES
 source
 1. 5450
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 538. .2535
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD44451.1"
 /db_xref="GI:22316985"
 /translation="MAHEMIGTQIVTERVALLESGTEKYLIDSPPEVYNTSHLE
 AININCSKLMKRLQODKVLITELIHSAGHKVDIOSOVYVYDDSDVASLSDC
 FLTVLGLKERSFNSVHLGAFSEFRCPLCEGSTIVPTICISQPLPVANIGPT
 RILPMLVGCORIVANKELIQONIGVYNASTCCRPDIPESHLRVPVDSFEK
 ILPLDKSVDFIIEKAKNGCVIVHCKIKQTGASGSKLILHLEKPEFVPVASEGQKSE
 TRLPSPCADSAESAQORVHPASVPSVQPSILESPVQALSGHLISADLEED
 SNKIKRSPSIDISVSASMAASLHGFSSSEDALEYKPSITLDTGKTKLCOFSPVDE
 LSQTPRTSPDKKEASIPKLIQTARPSDSQSKRLSHRTVRSSTGTORSLPLRHSQS
 VEDVHTSPFLEGLSTSQOHLTKSAGLGLGMSHDLIAPOTSPSLTSSWTFATSSHR
 YSASAIYGGASVAYSQSLPTCGDQVYSVRRROKPSRARSMSWESPFERKQFK
 RRSQCMFEGSINSENRRELGKVSQSPSGSMELIIVS"

ORIGIN

Alignment Scores:

Pred. No.: 1 4e-204
 Score: 3418.00 Length: 5450
 Percent Similarity: 100.00% Matches: 665
 Best Local Similarity: 100.00% Conservative: 0
 Query Match: 100.00% Mismatches: 0
 Gaps: 0

US-10-029-345a-109 (1-665) x AX482439 (1-5450)

QY 1 MetAlaHiGluMeClleGIyThrgInIleValThrgIuArgLeuValAlaleuLeuGlu 20
 Db 538 ATGGCCCAAGATGATTGGAACCAATGTTTCTAGAGAGTTGGTGGCTGCTGGGA 597
 QY 21 SerGIyThrgIuLyValleuLeuIleAsnSerArgProPheValGluTyraenThrSer 40
 Db 538 AGTGAACGGAAGAGCTCTCTTAATGATAGCCGGCATTGTTGGAAATACATATCATCC 657
 QY 41 HisIleuGluAlaIleAsnIleAsnCySerIleuMetIleuValArgIleuGln 60
 Db 658 CACATTGGAAGCATTATATCATCTGCCAGCTTATGAGCGAAGAGTTGGCAACG 717
 QY 61 AspIleuValleuIleThrgIuLeuIleGlnHisSerAlaIleuValIleuValIleuVal 80
 Db 718 GACAAAGTTTAATTAAGAGCTCATTCAGCATTCAGCAAAACATTAAGTTGACATTGAT 777
 QY 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIleValIleVal 100
 Db 778 TGGAGTGAAGAGTTGATTTACGATCAAGCTCCCAAGATTGCTCTCTCTCTTCA 837
 QY 101 AspCysPheLeuThrValleuLeuGluIleuGluIleuValIleuValIleuValIleuVal 120

Db 838 GACTGTTTCTACCTACTTCTGCGTAACTGGAAGACCTTCACTGTTCACTG 897
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyIleuCysGluGlyIleuSer 140
 Db 898 CTTCAGAGTGGGTTTGGAGTTTCTCGTGTGTTTCCCTGCGCTCTCTGAGGAAATCC 957
 QY 141 ThrIleuValProThrCysIleSerGlnProCysLeuProValAlaIleuIleGlyProThr 160
 Db 958 ACTCTAGTCCCTACTGATTCTCAGCTTCTGCTTACTCTGTCCAAACATTGGGCCAAC 1017
 QY 161 ArgIleuProAsnIleuTyraenGlyCysGlnArgAspValleuAsnIleuGluIleu 180
 Db 1018 CCAATTTCTCCCAATCTTATCTGCTGCGCAGCGAGATGCTCTCAACAGAGCTGATA 1077
 QY 181 GlnGluAsnGlyIleGlyIleValIleuAsnAlaSerTyraenCysProIleuProAspPhe 200
 Db 1078 CAGCAGATGGGATTGTTATGTTTAAATGACAGCTATATCTGCTCAAGCCTGACTTT 1137
 QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIleuValIleu 220
 Db 1138 ATCCCGAGTCTCATTTCTCGCTGCTGCTGGAATGACGCTTTTGTGAAGAAATTTG 1197
 QY 221 ProIleuAspIleuSerValAspPheIleuGluValIleuValIleuValIleuValIleu 240
 Db 1198 CCGTGGTGGACAACATCAGTAAATTTCTGAGAAAGCAAGCCTCCAAATGAGTGTGTT 1257
 QY 241 LeuValHisCysIleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyraenIleMet 260
 Db 1258 CTAGTGACCTGTTAGCTGGAGTCTCCGCTCCGCAATGCTATGCTTACATATG 1317
 QY 261 LysArgMetAspMetSerLeuAspGluIleValTyraenPheValIleuGluValIleuValIleu 280
 Db 1318 AAGAGGATGACATGCTTTGATGATTAACCTTACAGATTTGGAAGAAAGAAAGAAAGAAAGAAAG 1377
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnIleuLeuAspTyraenGluIleuValIleuValIleu 300
 Db 1378 ATATCTCCAAACTTCAATTTCTGCGCAACTCTGAGATATGAGAAAGAAAGTTAAGAAC 1437
 QY 301 GlnThrgIleAlaSerGlyProIleuSerIleuValleuIleuIleuIleuIleuIleuIleu 320
 Db 1438 CAGACTGAGCATCAGAGCCAAAGAGCAAACTAGCTGCTGAGTGAAGCAACAT 1497
 QY 321 GluProValProAlaValIleuGluGlyGlnIleuValIleuValIleuValIleuValIleu 340
 Db 1498 GAACCTGCTCTGCTGCTGCTGAGAGGTTGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557
 QY 341 CysAlaAspSerAlaThrSerGluAlaIleuGluIleuValIleuValIleuValIleuValIleu 360
 Db 1558 TGTGCCGACTGCTGCTGCTGAGAGGAGAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617
 QY 361 ProSerValProSerValGlnProSerIleuLeuIleuValIleuValIleuValIleuValIleu 380
 Db 1618 CCGAGCTGCTGCTGCTGAG 1677
 QY 381 SerGlyIleuHisIleuSerAlaAspArgIleuGluIleuValIleuValIleuValIleuValIleu 400
 Db 1678 AGTGGGCTGCACTGCTGCTGAG 1737
 QY 401 SerLeuAspIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 420
 Db 1738 TCTCTGAGATCAAAATCAAGTTTCAATTCATTCAGCAGCATGAGAGATCTTCAATGAGCTTC 1797
 QY 421 SerSerSerGluAspAlaIleuGluIleuValIleuValIleuValIleuValIleuValIleuVal 440
 Db 1798 TCCTCATCAAGAAAGATGCTTGAATATCAAAACCTTCACTACTGTGATGGAGCAAC 1857
 QY 441 LysIleuCysGlnIleuSerProValGlnIleuSerGluGlnIleuValIleuValIleuValIleuVal 460
 Db 1858 AAGCTATCCAGTTCTCTCTGCTGAGAGATATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1917
 QY 461 AspIleuGluIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 480

QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIlySerGluYlySerProThr 280
 Db 1318 AAGAGATGAGCATCTTTAGATGAAAGCTTACAGATTGTGAAGAAAAAAGAACTTACT 1377
 QY 281 IISerProAsnPheAsnPheLeuGlyGlnLeuAspPheGlnIlySerIlySerIlySer 300
 Db 1378 ATATCTCCAACTCAATTTTCTGTGGCCAACTCTGAGACTTATGAAAGAAAGATTAAAGAC 1437
 QY 301 GlnThrGlyAlaSerGlyProIlySerIlySerIlySerLeuIlySerGlnIlySerProAsn 320
 Db 1438 CAGACTGAGCATCAGGAGCCAAAGAGCAACTCAGCTGCTGCACTGGAGAGCAACT 1497
 QY 321 GluProValProAlaValSerGluGlyGlnIlySerGlnIlySerProLeuSerProPro 340
 Db 1498 GAACCTGTCCCTGCTCTCAGAGGGTGAAGCAAAAAGGAGAGCCCTCAGTCCACC 1557
 QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValIleProAlaSerVal 360
 Db 1558 TGTGCCACTCTCTACTCTCAGAGGAGCAGAGCAAGAGCCCGTGCATCCCGCAGCGTG 1617
 QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 Db 1618 CCCAGGCTGCCAGCTGAGCGCTGCTTAAAGAGCACGCCGCTGGTAAACAGCGCTC 1677
 QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIlySerIlySerPhe 400
 Db 1678 AGTGGCTGCACCTGTCTCCGAGACAGGCTGAGAGACAGCAATAGCTCAAGCGTCTCTC 1737
 QY 401 SerLeuAspIleIlySerValSerIlySerAlaSerMetAlaIleSerLeuHisGlyPhe 420
 Db 1738 TCTCTGGATTCMAATCAGTTTCATATTACGCCAGCATGCGACATCTTCAATGGCTTC 1797
 QY 421 SerSerSerGluAspAlaLeuGluTyrTyrIlySerProSerThrLeuAspGlyThrAsn 440
 Db 1798 TCCTCATCAGAAAGATCTTTGGAAATCTACAAACCTTCCACTCTGATGGAGCCAC 1857
 QY 441 LysLeuGlyGlnPheSerProValGlnIlySerGlnIlySerProIlySerPro 460
 Db 1858 AAGCTATGCCAGTCTCTCCCTGTTCAAGAACTATCGAGACAGATCCGAAACCAAGTCT 1917
 QY 461 AspIlySerGluAlaSerIleProIlySerLeuGlnIlySerAlaArgProSerAspSerGln 480
 Db 1918 GATAGAGAGAGACCCAGCATCCCAAGAAAGCTGCAAGCCGACGCTTCAAGACGCCAG 1977
 QY 481 SerIlyArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 Db 1978 AGCAGAGGATTCATTCGCTCAGAAACAGACAGATGGCACCCGCAAGAGTCTCTTTTA 2037
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnIlySerIlySerPheLeuPheGly 520
 Db 2038 TCTCCACTGCATCGAAGTGGAGGCTGAGAGCAATTACCAACACCACTTCTTTTCCGC 2097
 QY 521 LeuSerThrSerGlnGlnHisLeuThrIlySerSerAlaGlyLeuGlyLeuIlySerGlyTyrPhe 540
 Db 2098 CTTTCACACAGCAGACACACCTCAAGAACTCTGCTGCGCTTGAAGGCTGAC 2157
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerIlyIlyPheAla 560
 Db 2158 TCGGATATCTTGCCCCCAAGCTCTACCCCTTCCCTGACAGACGCTGATTTTGGC 2217
 QY 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyIlySerAlaSerIlySer 580
 Db 2218 ACGAGTCTTCACTTCTACTCTGCTCAGCCATCTACGAGGAGGAGCTCCAGTTACTT 2277
 QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValIlySerValArgArgArg 600
 Db 2278 GCTTACAGCTGACGACGCTGCCACTTGGGAGCAACAGTCTATTCTGTGGAGGCGG 2337
 QY 601 GlnIlySerSerAspArgAlaAspSerArgSerTyrPheIleGluGlnSerProPheIly 620
 Db 2338 CAGAGCCAAAGTACAGAGCTGACTCGCGCGGAGCTGSCATGAAGAGAGCCCTTTGAA 2397

QY 621 LysGlnPheIlySerArgSerCysGlnMetGluPheGlyGlnSerIleMetSerGluAsn 640
 Db 2398 AAGCAGTTTAAACGCAAGAGCTGCAAAATGGAAATTTGAGAGACATCATCTCAGAGAC 2457
 QY 641 ArgSerArgGlnGluLeuGlyIlySerValGlySerGlnSerSerPheSerGlySerMetGlu 660
 Db 2458 AGGTCAAGGAGAGACTGGGAGAAAGTGGGAGTCAAGTCTGCTTTTGGGACAGATGAA 2517
 QY 661 IleIleGluValSer 665
 Db 2518 ATCATTTAGAGTCTCC 2532
 RESULT 3
 AX260342
 LOCUS
 DEFINITION Sequence 3 from Patent WO0173059.
 ACCESSION AX260342
 VERSION AX260342.1 GI:16509305
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Meyers, R. A.
 3692 and 2117: dual specificity phosphatase molecules and uses
 therefor
 Patent: WO 0173059-A 3 04-OCT-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 JOURNAL
 TITLE
 AUTHORS
 SOURCE
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,95e-204 Length: 1998
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-345A-109 (1-665) x AX260342 (1-1998)
 QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 Db 1 ATGGCCCATGAGATGATGGAACTCAAAATTGTACTGAGAGGTTGGCTGCTGGAA 60
 QY 21 SerGlyThrGluIlySerValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 Db 61 AGTGAACGGAAGAAAGTCTGCTTAAATGATGACCGGCAATTTGGATTAACAATACATCC 120
 QY 41 HisIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 60
 Db 121 CACATTTTGAAGCCATTAAATCAACTGCTCAAGCTTATGAAGCGAAGGTTCAACAG 180
 QY 61 AspIlyValIleuIleThrGluLeuIleGlnHisSerAlaIlySerIlySerValIleAsp 80
 Db 181 GACAAAGTTTAAATTAACAGAGCTCATTCACGATTCACGAAACCTAAGTTGACATTGAT 240
 QY 81 CysSerGlnIlySerValIleValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 Db 241 TCGAGTCAGAAAGTGTAGTTTACGATCAAACTCCCAAAATGTGCTCTCTCTTCA 300
 QY 101 AspCysPheLeuThrValLeuLeuGlyIlySerGlnIlySerPheAsnSerValHisLeu 120
 Db 301 GACTGTTTTCATCTGATCTTCTGGATTAAGTGAAGAGCTTCAACTGTTCACTG 360
 QY 121 LeuAlaGlyIlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlnIlySerSer 140
 Db 361 CTGAGGTGGGTTTGTGAGTTCTCTGTTTCTCGGCTCTGTGAGAGAAATTC 420

QY 141 ThrLeuValProthrCysIIleSerGlnProCysLeuProValAlaAsnIIleGlyProThr 160
 Db 421 ACTGTAGTCCCTACCTGATTTCTCAGCTGTGCTTACTGTGCAACATTTGGCCAAACC 480
 QY 161 ArgIIleLeuProAsnLeuTyrLeuGlyCysGlnAspValIleuAsnIleuLeu 180
 Db 481 CGAATTCCTCCAACTCTTATCTTGGCTGCAGAGATGCTCTCAACAAGAGAGCTGATG 540
 QY 181 GlnGlnAsnGlyIIleGlyTyrValLeuAsnIleSerTyrCysProIleProAspPhe 200
 Db 541 CACAGAGATGGGATGGTGTATGTGTTAAATGACAGCAATACCTGTCCAAAGCCTGACTT 600
 QY 201 IIleProGlnSerIleAspLeuArgValProValAspAspSerPheCysGlyIleu 220
 Db 601 ATCCCGAGCTCAATTCCTGCGGTGCTGTGAATACAGACTTTGTGTGAAAAATTTTG 660
 QY 221 ProTyrLeuAspIleSerValIleAspPheIleGlyIleValIleAsnIleSerAsnGlyCysVal 240
 Db 661 CCGTGTGTGACAAATAGTATGATTTTCAATTGAGAAACAAACCTCCATGATGTGTT 720
 QY 241 LeuValIleCysLeuAlaGlyIIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 Db 721 CTAGTGCACTGTTAGCTGGGATCTCCGCTCCGCAACATCCCTATCGCTCAATCATG 780
 QY 261 LysArgMetAspMetSerLeuAspGlyAlaTyrArgPheValIleGlyIleAspArgProThr 280
 Db 781 AAGAAGATGAGACATGCTTTAGATGAAAGCTTACAGATTGTGAAAGAAAAAAGACTTACT 840
 QY 281 IIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyIleValIleValAsn 300
 Db 841 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTAAGAAC 900
 QY 301 GlnThrGlyAlaSerGlyProIleSerIleValIleuIleuIleuIleuIleuIleuIleu 320
 Db 901 CAGACTGAGACATCAGGGCCAAAGAGCAACTCAAGCTGTGACCTGGAGAGCCAAAT 960
 QY 321 GlnProValProIleValSerGlyGlyGlyIleValSerGlyIleuProIleuSerProPro 340
 Db 961 GAACCTGTCTCTCTGTGAGAGGGTGAACAAGAAAGCAAGAGCCCTCAATCCACCC 1020
 QY 341 CysAlaAspSerAlaThrSerGlyAlaAlaGlyIleValIleuProValIleProIleVal 360
 Db 1021 TGTGCGACCTCTGCTACTCAGAGGAGCAGAGCAAAAGCCCGCTGCATCCCGCAGCGTG 1080
 QY 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
 Db 1081 CCCAGCGTCCAGCGCAGCGCGCTGCTGTAGAGGACAGCCCGCTGTACAGCGCTC 1140
 QY 381 SerGlyLeuIleIleuSerAlaAspArgLeuGlnAspSerAlaIleuValArgSerPhe 400
 Db 1141 AGTGGGTGACCTGTCTCGAGAGAGCTGGAAGACAGCAATAGCTCAAGCCCTTCTTC 1200
 QY 401 SerLeuAspIleIleuSerValSerTyrSerAlaSerMetAlaIleSerLeuIleGlyPhe 420
 Db 1201 TCTCTGATATCAATATCATATTCATTCAGCAGCATGCGACATCTTACATGGCTTC 1260
 QY 421 SerSerSerGlnAspAlaLeuGlyIleTyrIleProSerThrThrLeuAspGlyIleAsn 440
 Db 1261 TCTCTCATCAGAGATGCTTGAATACTCAAACTTCCTCACTGATGAGGAGCAAC 1320
 QY 441 LysLeuCysGlnIleSerProValGlnIleuSerGlyGlnIleProGlyIleThrSerPro 460
 Db 1321 AACCTATGCAAGTTCCTCTGTTCAGGAACATTCGAGAGAGACTCCCGAAACAGAGCTC 1380
 QY 461 AspLysGlyGlnIleSerIleProIleValLeuGlnIleThrAlaArgProSerAspSerGln 480
 Db 1381 GATTAAGAGAGAGAGCAGCATCCCAAGAGACTGCAAGCCCGAGCTTACAGAGCCAG 1440
 QY 481 SerIleArgLeuIleSerValArgThrSerSerSerGlyIleThrAlaGlnArgSerLeu 500
 Db 1441 AGCAAGAGATGATTCGTCTGAGAACAGAGAGAGAGTGGCACCCCGCAGAGTCCCTTTTA 1500

QY 501 SerProLeuIleArgSerGlySerValGluAspAsnTyrIleThrSerPheLeuPheGly 520
 Db 1501 TCTCCATGCTGATCGAAGTGGAGCGTGGAGAGCAATATTACACACAGCTTCTTCGGC 1560
 QY 521 LeuSerThrSerGlnIleIleuThrIleValSerAlaGlyIleuGlyIleuValIlePhe 540
 Db 1561 CTTTCACAGCAGCAGACAGCACTCAGAAATGCTGTGCGCTGGGCTTAAGGCGCTGAC 1620
 QY 541 SerAspIleLeuAlaProGlnIleThrSerThrProSerLeuThrSerSerTyrPheAla 560
 Db 1621 TCGGATATCTGTGCCCCCAGAGCTTACCTCCCTTCCACAGCAGCTGTATTTTGCC 1680
 QY 561 ThrGlnSerSerIlePheTyrSerAlaSerAlaIleTyrGlyIleSerAlaSerTyrSer 580
 Db 1681 ACAGATCTCTACACTTCTACTCTGCTTACGCAATTCAGAGAGAGCTGCAATTTCTCT 1740
 QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
 Db 1741 GCCTACAGCTGACAGCCAGCTGCCCACTTCCGAGACCAAGCTATTTCTGTGCGCAGCG 1800
 QY 601 GlnIleProSerAspArgAlaAspSerArgArgSerTyrPheIleGlnIleuSerProPheGly 620
 Db 1801 CAGAACCAAGTACAGAGACTGCTGCGCGAGCTGCGATGAGAGAGAGAGAGAGAGAGAG 1860
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlyIleSerIleMetSerGluAsn 640
 Db 1861 AAGCAATTTAAACGCAAGAGCTGCCAAATGGAATTTGGAGAGAGCATCTGTACAGAAC 1920
 QY 641 ArgSerArgGlnIleuGlyIleValIleGlySerGlnSerSerPheSerGlySerMetGly 660
 Db 1921 AGGTACAGGGAAGAGCTGGGGAAGAGGAGGAGCACTAGCTTTTCGGGAGAGATGAA 1980
 QY 661 IleIleGluValSer 665
 Db 1981 ATCATTTAGAGTCTCC 1995

RESULT 4
 AX180875 LOCUS AX180875 2732 bp DNA linear PAT 06-AUG-2001
 DEFINITION Sequence 2 from Patent WO0146394.
 ACCESSION AX180875
 VERSION AX180875.1 GI:15132703
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 Plozman, G. D., Martinez, R., Whyte, D., Manning, G., Sudarshan, S.,
 Hill, R. J. and Flanagan, P.
 Mammalian protein phosphatases
 Patent: WO 0146394-A 2 28-JUN-2001;
 Sugen, Inc. (US)
 FEATURES
 source 1..2732
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3 02e-204 Length: 2732
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-345A-109 (1-665) x AX180875 (1-2732)
 QY 1 MetAlaHleGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 Db 538 ATGGCCCATGATGATTTGAATCTCAATTTTACTGAGAGAGTGGTGGCTGTGCTGAA 597

TITLE Identification of a dual specificity phosphatase: dusp-10
JOURNAL Patent: WO 0177340-A, 1 18-Oct-2001;
MERCK PATENT GmbH (DE)

FEATURES

Source

Location/Qualifiers

1..3059

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

CDS

127..2124

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD10545.1"

/db_xref="GI:16605916"

/db_xref="REFSEQ:CAD10545"

/translation="MAHEMIGDTITVERLVALLSESTERYLLIDSPFVETNTSHLE
ALININCSKLMKRLQODKYLITELIGHSAHKYDIDSKVVDVDSDDVALSDC
FLIVLGLKLEKSPNSVHLAAGFAEPRGCTVPCISQPLPVANIGPT
RLIPNLVLCORVDINKELMQONGIGVLANASNTCKPPIPESHFLRPVDSFCEK
ILPMLDKSVDFIEKAKNSGCVLHCLAGISRAATLAIYIMKRMMSIDEAVRPFKE
KPTTISPNRFLGQLDYERKIKNOVGASGRSKLLHLHEKNEPVPSGGQKSE
TPSPPCADNATSEAGORPVHPASVPSPVQPSLLEDSPLVALSGHLSADRLSD
SNLKRKSFLLDKSVASVSMASLHGFSSBEDALYKRPSTLIDGTNKLQPSVQK
LSEQETSPDKKEASIPKLTQAPRSDSQKRLHSVRSSGTAQRSLISPLRSGS
VEDNHTSFLFGLSTSOHLTKSAGLGLQMHSDILAPOSTPSLTSWYPATESHF
VSASAIYGGSAVSAYSACSLPTPCSDOVYSVRRROKPSDRSRMHESPPEKQFK
RNSCOMFEGESIMSENRREBELKVSQSSFSQSMETIEYS"

ORIGIN

Alignment Scores:

Pred. No.: 3, 53e-204 Length: 3059
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 0
Query Match: 99.65% Indels: 0
Gaps: 0

US-10-029-345A-109 (1-665) x AK278461 (1-3059)

QY 1 MetAlHiEGlUeCtIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 127 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 186
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 187 AGTGAACGGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 246
QY 41 HisIleLeuGluValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 247 CACATTTTGGAAACCATTAATCACTGCTCCAAAGCTTATGAAGGAGGATGGCAACAG 306
QY 61 AspLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 307 GACAAAGTGTATATACAGAGCTCATCCAGCATTCAGCGAAACATGAAGGTGACATTGAT 366
QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 367 TGCAGTCAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
DB 427 GACTGTTTCTCACTGACTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACTG 486
QY 121 LeuAlaGlyLysPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlnGlyLysSer 140
DB 487 CTTCGAGGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
QY 141 ThrLeuValProThrCysAlaSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 547 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGCTTACCTGTTCGCAACATTTGGGCCAAC 606
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
DB 607 CGAATCTTCCCATCTTTATCTTGGCTGCGACAGAGATGTCCTCAACAAAGAGCTGATG 666

QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
DB 667 CAGCGAATGAGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 726
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
DB 727 ATCCCGAGATCTCATTTCTGGGTGCTGTGGAATGACAGCTTTTGTGAGAAATTTTG 786
QY 221 ProThrLeuAspLysSerValAspPheIleGlyLysValAlaSerLeuGlyCysVal 240
DB 787 CCGTGGTGTGAACAAATCAGTATGATTCATTTAGGAAAGCAAAAGCCCTCAATGATGTGTT 846
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
DB 847 CTAGTGCATGTTTACTGGGATCTCCCGCTCCGACACATGCTATCGCTACATCAAG 906
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
DB 907 AAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTGTGAAGAAAAAGACTACT 966
QY 281 IleSerProAsnPheAsnPheLeuGlyLysLeuLeuAspTyrGluLysLysIleLysAsn 300
DB 967 ATATCTCCAACTTCATTTCTGGGCCAACTCTGCACTATGAGAAAGATTAAGAAG 1026
QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGluLysProAsn 320
DB 1027 CAGACTGAGGACATCAGAGGCGCAAGAGCAAACTCAAGCTCTGACCTGAGAGGCCAAAT 1086
QY 321 GluProValProAlaValSerGlnGlyLysLeuSerGlyThrProLeuSerProPro 340
DB 1087 GAACCTGTCTCTGCTGTCTCAGAGGATGACAGAAAGAGAGAGCGCTCATGTCACACC 1146
QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyLysGlnArgProValHisProAlaSerVal 360
DB 1147 TGTGCCGACTGCTACCTCAGAGGACAGAGCAAAAGCCGCTGATCCCGCACGCTG 1206
QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
DB 1207 CCAGGCTCCCAAGCGCTGAGCGCTGTGTTAGAGAGACAGCCGCTGTGACAGGCGCTC 1266
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
DB 1267 AGTGGCTGACCTGTGCTCCGACAGACAGCTGGAAGACAGCAATAGCTCAAGCTTCTTC 1326
QY 401 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
DB 1327 TCTCTGATATCAATCAATCAGTTTCATATTCAGCCAGCATGGCAGCATCTTACATGCTTC 1386
QY 421 SerSerSerGluAspAlaLeuGlnTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
DB 1387 TCCTCATCGAAGATGCTTTGAAATTCACAACTTTCATCTGTGATGGAGACCAAC 1446
QY 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
DB 1447 AAGCTATGCCAGTTCTCCCTGTTTCAAGAACTATGCGAGACACTCCCAAAACAATCTCT 1506
QY 461 AspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
DB 1507 GATTAAGAGAGAGAGCAGCATCCCAAGAGCTGACAGCCGACGCTTCAAGACAGCCAG 1566
QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
DB 1567 AGCAAGCGATTCATTCGCTCGAGAACACACAGACAGAGGAGCCGCCAGAGAGCTCTTTTA 1626
QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
DB 1627 TCTTCATCTGATCGAAGTGGAGCGTGGAGAGCAATTTACACACAGCTTCTCTTTGGGC 1686
QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysLeuGlyTyrPhe 540
DB 1687 CTTTCCACAGCAGCAGACAGCACTCAAGAGTGTGCTGGGCTTAAAGGAGCTGGAC 1746

QY	541	SetrApIleuAlaProGIntnrSerThrProSerLeuThrSerSerTrpYrPheAla	560
Db	1747	TCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACGAGAGTGATATTGGCC	1801
QY	561	ThrGluSerSerHisPheTrpSerAlaSerAlaIleYrGlyGlySerAlaSerYrSer	580
Db	1807	ACAGAGCTCTCACACTTCTTACTCTGCTTACGCACTTACGAGGACAGTGCACATCTACTCT	1866
QY	581	AlaYrSerCysSerGlnLeuProThrCysGlyApbGlnValYrSerValaGaaGaaG	600
Db	1667	GCTTACAGCTGACGACGAGCTGCCACTTGGCCGAGACCAAGCTATTTCTGTGGCGAAGCG	1926
QY	601	GlnYrProSerAspArgAlaAspSerArgYrSerTrpHisGluGluSerProPheGlu	620
Db	1927	CAGAACCCAAAGTACAGAGCTGACCTCCGGCGGAGAGCTGGCATGAMAGAGCCCTTTGAA	1986
QY	621	LYeGlnPhePheArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluGln	640
Db	1987	AAGCACTTAAAGGACGAAGCTGCCAAATGGAAATTGGAGAGGACATATGTCAAGAAC	2046
QY	641	ArgSerArgGluGluLeuGlyYrValaGlySerGlnSerSerPheSerGlySerMetGlu	660
Db	2047	AGGTACCGGAAAGAGCTGGGAAAGTGGCGAGTCACTTACTTTCGGCGACATGAA	2106
QY	661	IleIleGluValSer	665
Db	2107	ATCATTTAGTCTCTCC	2121
RESULT 6			
AX405700			
LOCUS	AX405700	3104 bp	DNA
DEFINITION	Sequence 115 from Patent WO0222660.	linear	PAT 14-JUN-2002
ACCESSION	AX405700		
VERSION	AX405700.1	GI:21438839	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE			
AUTHORS	Tiang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,		
JOURNAL	Xue, A.J., Yang, Y., Wehrman, T., and Demanac, K.T.		
	Novel nucleic acids and polypeptides		
	Patent: WO 0222660-A 115 21-MAR-2002;		
	HYSEQ, INC. (US)		
FEATURES			
source	Location/Qualifiers		
	1..3104		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	159..2156		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD34807.1"		
	/db_xref="GI:21438840"		
	/db_xref="REMBEMBL:CAD34807"		
	/translation="MAHEMIGTQIVTERLVLLSGTEFKLITLIDSPFEVYNTSHILE		
	ALINTNSCKLMKRRLQDDKVLITELIOHSAKKHVIDDSQYVYTDSSQDVASLSDC		
	FLTVLGLKLEKSFNSVHLAAGFAEFSFCPLCEGSTVPTCTIOPCLPAVANGT		
	RILPNTLYLGGDRDVLNKLKELMGQNGVLYNLSNTCPEDTIPESHRLVNPVDFEET		
	ILPMLDKSVDFTEKAKASGNCVLYCLAGLISRTAIAVIMRMMSIDEAYRVKE		
	KRTPTIPNPNFLGOLLIDYEKTKIKNOTGASGSKLKLILHEKNEPEFPAVSBGQKSE		
	SNLKTSKPSLIDKISYISASMAASLHPSSEEDALEYRKSTPLDGTNLSGLHLSADBLE		
	LSQPTPEPDEKEASIPRKQTARPSIQSGSKRLHSVTSSSGTAQRSLISPLHRSGS		
	VEDNHTSFLFGLSTSQHLTKLSAGLGLKGMHSPILIAPOSTPSLITSMYPTIESHP		
	YKSAWYGGSSASVAYSCSOLPTCGDQVYVTRERKRPSPRADRSRMSWHESPFEKQFK		
	RRSQWEPEGESIMSHNRREBELGVGSSPFSMSMELIYS"		
ORIGIN			
Alignment Scores:			
Prd. No.:	3 61e-204	Length:	3104
Score:	3406.00	Matches:	663

	Percent Similarity:	99.85%	Conservative:	1
	Best Local Similarity:	99.70%	Mismatches:	1
	Query Match:	99.65%	Indels:	0
DB:		6	Gaps:	0
US-10-029-345A-109 (1-665) x AX405700 (1-3104)				
QY	1	MetaIaHISgluMetIIEglYThrGlnIleValThGluArgLeuValAlaLeuLeuGlu		20
DB	159	ATGGCCCATGATGATGATTTGGAACCTCAAAATTTGATTACGAGAAGTGGTGCTCGTGGGA		21
QY	21	SerGIYThrGluIuYValLeuLeuIleAspSerArgProPheValGluTrzAsmThSer		40
DB	219	AGTGAACCGAAAAAGGTGCTAAATTGATACCGGGCCATTGTGGAAATACAAATCATTC		27
QY	41	HisIleuGluIuAlaIleAsnIleAsnCySerIYsLeuMetLYsArgArgLeuGlnGln		60
DB	279	CACATTTTGGAAACCATTAATATCAACGTCTCCAAAGCTTAAGAAGCGAAGGTTCACACAG		333
QY	61	AspLYsValLeuIleThrGluLeuIleGlnHisSerAlaLYSHsLYsValAspIleAsp		80
DB	339	GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGAACATTGAT		398
QY	81	CysSerGlnLYsValValIYTrzAspGlnSerGlnAspValAlaSerLeuSerSer		100
DB	399	TGCAGTCAGAAAGTTGATGATTTACGATCAAGCTCCAAAGATGTTGCTCTCTCTTCA		458
QY	101	AspCysPheLeuThrValLeuLeuGluLYsLeuGluIuYsSerPheAsnSerValIHisLeu		120
DB	459	GACTGTTTCTCACTGATCTTCTGGGAAACCTGGAGAAGAGCTTCAACTCTGTTCACTCG		518
QY	121	LeuAlaGlyIYpheaIagluPheSerArgCysPheProGluYleuCYsGluGluLYsSer		140
DB	519	CTTCAGAGTGGTTCGTGAGTTCTCTCGTTGTTTCCCTGCGCTCTGTAAGAGAAATTC		578
QY	141	ThreValIProThrCysIleSerGlnProCysLeuProValAlaAsnIIEglYProThr		160
DB	579	ACTGTATCTCCATCTGACATTTCTCAAGCTTGCTTACTGTTGCGCAACATTGGGCCAAC		638
QY	161	ArgIleLeuProAsnLeuTYrLeuGluCYsGlnArgAspValLeuAsnLYsGluLeuIle		180
DB	639	CGAATCTTCCCAATCTTATCTTGGCTCGACGAGATGTCCTCAACAAAGAGTGAATG		698
QY	181	GlnGlnAsnGlyIIEglYTYrValLeuAsnAlaSerTYrThrCysProLYsProAspPhe		200
DB	699	CACACAGAAAGGATGGTATGTATGTGTAATGCAACAATACCTGTCACAAACCTGACATT		758
QY	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLYsIleLeu		220
DB	759	ATCCCGAGATCTCAATTTCTCGTGTCGCTGATGATGACACTTTGTGAAAAATTTTG		818
QY	221	ProTYrLeuAspLYsSerValAspPheIleGluLYsAlaLYsAlaSerAsnGlyCYsVal		240
DB	819	CCGTGGTGGACAAATCGATGATTTCATTGAGAAACAAAGAGCTCCAAATGATGTGT		878
QY	241	IeuValHisCYsIleuAlaGlyIIEserArgSerAlaThrIleAlaIleAlaTYrIleMet		260
DB	879	CTAGTGCACTGTTTGTGCTGGGATCTCCCGCTCGGCCACCATCTATCGCTATACATCATG		938
QY	261	LYsArgMetAspMetSerLeuAspGluAlaTYrArgPheValLYsGluLYsArgProThr		280
DB	939	AAGAGGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAACAAAAAGACCTTACT		998
QY	281	IIEserProAsnPheAsnPheLeuGlyGlnLeuLeuAspTYrGluLYsLYsIleLYsAsn		300
DB	999	ATATCTCCCAATCTCAATTTTCTGGGCCAAACTCTCGGACCTATGACAAAGATTAAAGAC		1058
QY	301	GlnThrGlyAlaSerGlyProLYsSerLYsLeuLYsLeuLeuHisIleuGluLYsProAsn		320
DB	1059	CAGACTGGAGCAACAGAGGCCAAAGAGCAAACTCAAGCTCTGCACTCTGGAAGGCAAAAT		1118
QY	321	GluProValProAlaValSerGlnGlnGlyIIEglYProThr		1118

Db 1119 GAACCTGCTCCTGCTGCTCAGAGGTGAGACAGAAAGGAGAGCCCTCATGTCACCC 1178
 Qy 341 CyeAlaAapSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 Db 1179 TGTGCCACTCTGCTACTCAGAGGAGCAGAGCAAAAGGCCCGCTGATCCCGCAGGTG 1238
 Qy 361 ProSerValProSerValGlnProSerLeuGluAlaAapSerProLeuValGlnAlaLeu 380
 Db 1239 CCCAGCTGCCCGCCAGCTGACGCCCTCGTGTAGAGACAGCCGCTGTGACAGGCCCTC 1298
 Qy 381 SerGlyLeuHisLeuSerAlaAapArgLeuGluAapSerAlaHisLeuValArgSerPhe 400
 Db 1299 AGTGGCTGACCTGCTCCGACAGACAGCTGGAAGACATAGCTCAAGCGTTCCTTC 1358
 Qy 401 SerLeuAapLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 Db 1359 TCTCTGATATCAAAATCAGTTTCAATATTCAGCCAGCTGACAGATCTTACATGCTTC 1418
 Qy 421 SerSerSerGluAapAlaLeuGluTyrTyrLysProSerThrThrLeuAapGlyThrAsn 440
 Db 1419 TCCTCATCAAGAAATGCTTTGGATTAATACTCAAACTTCCACTCTGTGATGGAGCCAC 1478
 Qy 441 LysLeuGlnProSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
 Db 1479 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATCGAGAGACTCCGAAACAGTCTC 1538
 Qy 461 AspLysGluGluAlaSerLysProLysLeuGlnThrAlaArgProSerAapSerGln 480
 Db 1539 GATAGAGAGAGAGCAGCAGATCCCAAGAGCTGACAGCTCCAGGCTTCAAGACAGCAG 1598
 Qy 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 Db 1599 AGCAAGGATTCGATTCGTCAGAACAGACAGCAGTGCACCGCCAGAGTCCCTTTTA 1658
 Qy 501 SerProLeuHisArgSerGlySerValGluAapAsnTyrHisThrSerPheLeuPheGly 520
 Db 1659 TCTCACTGATCGAAGTGGAGGTGAGGAGCAATTAACACACAGCTTCCTTTGGC 1718
 Qy 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrHis 540
 Db 1719 CTTTCCACGAGCAGCAGCAGCAGCTTCAAGAGTCTGCTGGGCTTAAAGGCTGGCAC 1778
 Qy 541 SerAapLysLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrTyrPheAla 560
 Db 1779 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGCC 1838
 Qy 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer 580
 Db 1839 ACAGAGTCCCTCACACTTCTACTGCTGCTCAGCCATCTAGAGAGGAGTGCAGTTACTT 1898
 Qy 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAapGlnValTyrSerValArgArgArg 600
 Db 1899 GCTTACAGCTGACAGCAGCTGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGAA 1958
 Qy 601 GlnLysProSerAapArgAlaAapSerArgSerTyrHisGluGlnArgProPheGlu 620
 Db 1959 CAGAAAGCAAGTACAGAGCTGACTCCGGGAGCTGGGATGAGAGAGAGCCCTTTGAA 2018
 Qy 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyLysSerLysMetSerGluAsn 640
 Db 2019 AACAGATTAAAGCAGAAAGCTGCAAAATGGAAATTTGGAGAGCATCATGTCAAGAAAC 2078
 Qy 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 Db 2079 AGGTTCACGGGAAAGCTGGGAGAAAGTGGAGTCAAGTCTTCTTCCGCGACAGATGAA 2138
 Qy 661 IleIleGluValSer 665
 Db 2139 ATCATTCAGGTCTCC 2153
 RESULT 7
 AX441210
 LOCUS AX441210 3496 bp DNA linear PAT 28-JUN-2002

DEFINITION Sequence 1 from Patent WO0226997.
 ACCESSION AX441210
 VERSION AX441210.1 GI:21665766
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Lucche, R.M. and Wei, B.
 TITLE Dep-16 dual-specificity phosphatase
 JOURNAL Patent: WO 0226997-A 1 04-APR-2002;
 Ceptiv, Inc. (US)
 FEATURES
 source Location/Qualifiers
 1..3496
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.26e-204 Length: 3496
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: Gaps: 0
 US-10-029-345A-109 (1-665) x AX441210 (1-3496)
 Qy 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 Db 562 ATGGCCATGAGATGATGAGTGAATGACTCAATATGTTACTGAGAGGTGGTGGCTGCGGAA 621
 Qy 21 SerGlyThrGlnLysValLeuLeuLysAapSerArgProPheValGluTyrAaThrSer 40
 Db 622 AGTGAAGCGAAGAAATGCGTGAATTAATGATAGCGGCGCATTTGTGGAATACATACATCC 681
 Qy 41 HisIleLeuGluAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGln 60
 Db 682 CACATTTTGAAGCCATTAAATCAATGCTCCACATTAATGAAGGAGTTGCAACAG 741
 Qy 61 AspValLeuLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAapLysAap 80
 Db 742 GACAAAGTGTAAATTAAGAGCTCATCAGCATTCAGCGAAACATAAAGTTGACATTGAT 801
 Qy 81 CysSerGlnLysValValTyrAapGlnSerSerGlnAapValAlaSerLeuSerSer 100
 Db 802 TGCAGTCAGAGGTTTATGTTATGATCAAAAGCTCCCAAGATGTTCCTCTCTTCA 861
 Qy 101 AapCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
 Db 862 GACTGTTTCTCACTACTTCTGAGTAACTGGAAGAGAGTCAACTCTGTTCACTG 921
 Qy 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuGluGlyLysSer 140
 Db 922 CTTCAGAGGGGTTTCTAGATCTCTCGTGTTCCTCGGCTCTGTGAAGGAAATCC 981
 Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 Db 982 ACTCTAGTCCCTTACCGCATTTCTCAGCTTGCTTACCTGTTCCCAACATTTGGGCCAAC 1041
 Qy 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAapValLeuAsnLysGluLeu 180
 Db 1042 CGAATCTTCCCAATCTTATCTTGGCTCCAGCGAGATGCTCTCAACAGAGGCTGAAG 1101
 Qy 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysArgLysProAapPhe 200
 Db 1102 CAGCAAAATGGAGATGGTATGTATGTTAAATGCGAGCAATACCTGTTCAAAAGCTGACTTT 1161
 Qy 201 IleProGluSerHisPheLeuArgValProValAsnAapSerPheCysGluLysIleLeu 220
 Db 1162 ATCCCGAGTCTCATTTCTCTGGGTGCTGTGAATGAACAGCTTTTGTGAGAAATTTTG 1221

QY 221 ProTTrleuaplsysSerValaspPheIleglulysAlalysAserAenglyCysVal 240
Db 1222 CCGGGGTGGCAAAATCAGTAGATTTCATTGGAAGAAGCAAAGCTCCATTGAGATGTGT 1283
QY 241 LeuValHisCysLeuValaglYllsestrgsesrAlathrllealllealatlalatyrllemet 260
Db 1282 CTAGTGACACTTTTACTGTGGATCTTCGCCGCTCCGCACCATCTCCCTAACATCATCATG 1341
QY 261 LysArgMetAspMetSerleuaspgluAlatyrArpPheValLysglulysArpProthr 280
Db 1342 AANAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAGAAAGCTTACT 1401
QY 281 lIeSerProasnPheasnPhelenglGluleneuasDtyGluLylalsylleLysasn 300
Db 1402 AAATATCCAAACTTCATTTTCTGGCCAACTCTTGGACTTAATGAGAAGAAAGATTAAAGAC 1461
QY 301 GlNThrnglYAlaserglyProlysSerylLeuLylaleuLysleuHIsleuGlulysProasn 320
Db 1462 CAGACTGGAGCATCAGGGCCAAAGAGCAAATCMAAGCTCTGCACTTGGAAGACCMAAT 1521
QY 321 GluProvalProalavalsergluglyglInLysSerglutThrProleusSerPropro 340
Db 1522 GAACCTGTCCCTGTCTGTCTAGAGGGTGGCAGAAAGAGAGCCCCCTTAGTCCACC 1581
QY 341 CysAlaaspSeralatrRsrsgluAlaglYglInahgrProvalHisproalaserVal 360
Db 1582 TGTGGCCACTCTGTACTCAGAGGCGAGAGACAAGCCCCTGGCATCCGCCAGCGTGT 1641
QY 361 ProSerValProserValGlnProserleuengluaSerProleuValGlnAlaleu 380
Db 1642 CCCAGCCGGCCACGCGTGGAGCGCTGTGTAGAGACACGCCGCTGTGACAGGCGCTC 1701
QY 381 SerGlyleuHIsleuSerAlaaspArgleugluaspSerasnlyslenuylarsSerphe 400
Db 1702 AGTGGGTGGACCTGTCCGCACACAGCGCTGAACACAGCATTAAGCTCAAGCGTCTTC 1761
QY 401 SerleuasplleLysSerValSerylserAlaserMetAlaaserleuHIsglyPhe 420
Db 1762 TCTCTGGATTCMAATCMATTCATTAATCACCCAGCTGACACATCTTTACATGGCTTC 1821
QY 421 SerSerSerGluaspAlaleuenglulnryTyrlLysProSerThrThrleuaspglylThrasn 440
Db 1822 TCCTCATTCAGAAAGCTTTGGAATFACTAACAACTTCACATCTGTGATGGAGCCAAC 1881
QY 441 LysLeuCySGlnPheSerProvalGlnLysSerGlnLysSerGlnThrProgluthrSerPro 460
Db 1882 AAGCTATGGCAGTTCTCCCTGTTTCAGAAACATAAGGACAAATCCCGAAACACAGTCT 1941
QY 461 AsplyASgInglualaserlleprouLysleuGlnThralarPProSerAspSerGln 480
Db 1942 GATTAGAGAGAAAGCACATCCCAAGAAAGCTGCAGACGCCAGCCTTCAAGACCCAG 2001
QY 481 SerlysaArgleuHIsSerValArgThrRsrSerSerGlyThrAlaglnahgrSerleu 500
Db 2002 AGCAAGGCAATTCATTCGTGTCAAGACACACACACAGATGGACCGGCCAGAGGTCCCTTTTA 2061
QY 501 SerProleuHIsArgserGlySerValGluaspAsnyrHIsThrSerPheleuPhegly 520
Db 2062 TCTCACAATGCATGAAGTGGAGCGGTGAGACAAATTACCAACACACTTCCTTTGGGC 2121
QY 521 leuSerThrserGlnGlnHIsleuThrxlysserAlaglYleuuglyleuLysglYTPHIs 540
Db 2122 CTYTCCACACAGCAGACAGCACTTCAGAAAGTGTGTGCTGTGGCTTAAAGGGCTGGCAC 2181
QY 541 SerAspIlleuValaproGlnThrserThrProserleuThrserSerTrpyrPheAla 560
Db 2182 TCGGATATCTTGCCCCCAGACTTACCCCTTCCCTGACAGCAGCTGATATTGGC 2241
QY 561 ThrGlnSerSerHisPheTyrsenalaserAlaletyrglyserAlaserTyser 580
Db 2242 ACAGATCTCTCACACTTACTGTGCTCAGCCATCTTAAGAGGAGAGTCCAGATTACTCT 2301

QY 581 ALaTySerCySSerGInLeuPProHrCysGlyAspGlnValTyrSerValArgArg 600
Db 2302 GCCTACAGCTGCAGCCAGCTGCCACTTCGGAGACCAAGTCTATTCTGTGGCAGGGCG 2367
QY 601 GlnLysProSerAspArgAlaAspSerAlaGlySerTrpHisGlnGluSerProPheGlu 620
Db 2362 CAGAACCAAGTACACAGACTTACCTCGCGCGAGCTGGCATGAGAGAGCCCTTTGAA 2421
QY 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
Db 2422 AAGCACTTTTAAACGACGAAAGCTGCCAAATYGAAATTGGAGAGACATATATGCAGAGAC 2481
QY 641 ArgSerArgGlnGluIleGluGlyLysValGlySerGlnSerPheSerGlySerMetGlu 660
Db 2482 AGCTCAGCGGAAGACTCGGGAGAAAGTGGGACGTCAGTCTTTCGGGCGAGCATGAA 2541
QY 661 IleIleGlnValSer 665
Db 2542 ATCATTTAGAGTCTCC 2556

RESULT 8			
AB052156	AB052156		
LOCUS			
DEFINITION	Homo sapiens MKP-7	3521 bp	mRNA
ACCESSION	AB052156		linear
VERSION	AB052156.1	GI:13548676	PR1 18-OCT-2001
KEYWORDS			for MAPK phosphatase-7, complete cds.
SOURCE			
	Homo sapiens (human)		

SOURCE ORGANISM	ORGANISM
Homo sapiens (human)	Homo sapiens
Homo sapiens	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Homini	Mammalia; Eutheria; Primates; Catarrhini; Homini
REFERENCE 1	REFERENCE 1

ABBREVIATIONS MAPK-7, Shimizu, H., Watanabe, M. and Kikuchi, K. MKP-7, a novel mitogen-activated protein kinase phosphatase functions as a shuttling protein
JOURNAL J. Biol. Chem. 276 (42), 39002-39011 (2001)
JOURNAL MEDLINE 21486429

REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 3521)
Masuda, K., Shima, H. and Kikuchi, K.
Direct Submission
Submitted (05-DEC-2000) Kohsei Masuda

Department of Biochemical Oncology and Immunology,
National Institute of Health, Sapporo, Hokkaido 060-0815, Japan
(E-mail: koudai@nhi.ac.jp, Tel: 81-11-706-5536,
Fax: 81-11-707-6839)

Source	1. .3521
...	...

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/cell_line="U937"
1..3521
/gene="MKP-7"
564..2561
DSDS

```

```

/gene="MKP-7"
/codon_start=1
/product="MAPK phosphatase-7"
/protein_id="BAB40814.1"
/db_xref="GI:13548677"
/translacion="MAHEMIGTQIWLTERLVALLESTCEKVLIDSPPEVEYNTSHILAE
AININCSKIMKRRLOODKYLITGLIOHSAHKKVIDDSQGLVAVYQDSQDVASLSDD
FLVLYLGEKSPNSVHLLAGFAFSCFPGLECEGKSLVPTTCLISQPLPANIGPGL
LILBNLYLGQKRVYANKELMNOGIVGLNASTCRCPDPIESHFHRLVPAVDSCERK
ILPRDKSDVFLBKAAAGCNGVLLHCLAGLSRATTAIAVIMKRDMLDVAAYFVKE
KRPLTSPENFNGQLIDYEKTIKNQTSASPKKLLLEKNDPEVPAVSEGCQKEE
TLPSPCCASATSEAAQGPVHPASVPSPVSDPILLESPIVQLSGHLSARLDE
SNLKTSFELDIKSVYSASMAWALHGFSSSEALATYTPSTLLGNKLCQFSPVDE
LSBOTETEPIDKEEASIPKTKQTRPDSQSKLBSVIRTSSSGTAKORILSWPLHRSQS
VEANYHTSLRGLSTSOQHLTKSAGLCKMHSDILAPOTSPITSLTSWYATATSSHF
YASATVGSASATYSCSLPTCGQVSVVRKQSPDRADRSRSMHBSPEFKQK
RSCQWEPFESITSMENSRRELGQVQSOSFSGSMETIYVA"

```

ORIGIN

Alignment Scores:

Pred. No.: 4.3e-204 Length: 3521
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: 9 Gaps: 0

US-10-029-345a-109 (1-665) x AB052156 (1-3521)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
 DB 564 ATGGCCCATAGATGATTTGAACTCAATTTGTTACTGAGAGTTGGCTCTGCTGGAA 623
 QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyraenthSer 40
 DB 624 AGTGAACGGAAAGAGCTGCTAATTGATAGCCGGCATTTGGAAATACATATCCTC 683
 QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerLeuMetLeuValArgLeuGln 60
 DB 684 CACATTTTGGAAACCTTAATATCACTGCTCAAGCTTAAGAGCGAAGTTGCAACAG 743
 QY 61 AspLeuValLeuIleThrGluLeuIleGlnHisSerAlaValHisValAspIleAsp 80
 DB 744 GACAAAGTTAATACAGAGCTCATCCAGCATTCAGCGAAACATAGTTGACATTGAT 803
 QY 81 CysSerGlnValValValValTyraProGlnSerSerGlnAspValAlaSerLeuSer 100
 DB 804 TGCAGTCAGAGGTTGATTTGATTCAGATCAAGCTCCCAAGATTTGCTCTCTCTCA 863
 QY 101 AspCysPheLeuThrValLeuLeuGlyValLeuGluValSerPheAsnSerValHisLeu 120
 DB 864 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGCTTCACTGTTCACTG 923
 QY 121 LeuAlaGlyValPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlySer 140
 DB 924 CTTCAGAGTGGTTTGTCTGAGTTCTCTGTTGTTCCCTGCTGTAAGAAATCC 983
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProthr 160
 DB 984 ACTCTAGTCCCTACCTGCAATTTCTGAGCCTTGCTTACCTGTTGCAACATTTGGCCAA 1043
 QY 161 ArgIleLeuProAsnLeuTyraLeuGlyCysGlnArgAspValLeuAsnValGluLeu 180
 DB 1044 CGAATTCCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCTCAACAAAGAGCTGATG 1103
 QY 181 GlnGlnAsnGlyIleGlyTyraValLeuAsnAlaSerTyraCysProValProAspPhe 200
 DB 1104 CAGCAGATGGAGTGGTATGTTAATATGCAAGCATACCTGCTCAAAAGCTGACTTT 1163
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleLeu 220
 DB 1164 ATCCCGAGTCTATTTCCGCGTGGCCCTGTAATGACAGCTTTGTGTGAAGAAATTTTG 1223
 QY 221 ProTrpLeuAspIleValAspPheIleGluValAlaValAsnAsnGlyCysVal 240
 DB 1224 CCTGTGTGGACAAATCAGTAGATTTCAATGAGAAAGCAAAACCTCAATGATGTGTT 1283
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyraIleMet 260
 DB 1284 CTAGTGACTGTTAGTGGAGTCTCCCGCTCGGCACCAATGCTATGCTATCATCATG 1343
 QY 261 LysArgMetAspMetSerLeuAspGluAlaTyraArgPheValIleGluValArgProthr 280
 DB 1344 AAGAGATGAGCATGCTTTAGATGAAGCTTACAGATTTCTGAAGAAAGAAAGAAAGCTA 1403
 QY 281 IleSerProAsnPheAsnPheLeuGlyValLeuLeuAspTyraGluValIleValAsn 300
 DB 1404 ATATCTCCAAACTTCAATTTCTGGGCGCAATCTCTGAGCATATGAGAAAGATTAAGAAC 1463
 QY 301 GlnThrGlyAlaSerGlyProLysSerIleLeuLeuLeuHisLeuGluValSerProAsn 320

DB 1464 CAGACTGAGCATCAGGGCCAAAGACCAATCAAGCTCTGACCTGGAGAAACCAAT 1523
 QY 321 GluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProPro 340
 DB 1524 GAACCTGCTCCCTGCTGTCTCAGAGGGGTGACGAAAGAGAGCGCTCTGATTCACCC 1583
 QY 341 CysAlaAspSerAlaThrSerGluAlaIleGlyValArgProValHisProAlaSerVal 360
 DB 1584 TGTGCGACTGTCTACCTCAGAGGACGAGCAAGAAAGCCGTGTGATCCGCTCACGTTG 1643
 QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 DB 1644 CCGAGCTGCCAGCTGACGCTGAGCGTGTCTTAAAGAGACAGCCGCTGGTAAGAGGCTC 1703
 QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnValLeuValArgSerPhe 400
 DB 1704 AGTGGGCTCACCTGCTCCAGACAGCTGGAAGACAGCAATAGCTCAAGCTTCTTTC 1763
 QY 401 SerLeuAspIleLysSerValSerTyraSerAlaSerMetAlaSerLeuHisGlyPhe 420
 DB 1764 TCTTGATATCAATACAGTTTCAATTCAGCCAGCATGCGACATCTTACATGGCTTC 1823
 QY 421 SerSerSerGluAspAlaLeuGluTyraTyraLysProSerThrThrLeuAspGlyThrAsn 440
 DB 1824 TCTCATCAAGAGATGCTTTGGAATATCTCAAACTTCACTGATGGAGAACCAAC 1883
 QY 441 LysLeuCysGlnPheSerProValGlnLeuSerGluGlnThrProGluThrSerPro 460
 DB 1884 AAGCTATGCAATTTCTCCCTGTTGAGAACTATCGAACAAGCTCCCAAAACCAATCTCT 1943
 QY 461 AspLysGluGluAlaSerIleProLysValLeuGlnThrAlaArgProSerAspSerGln 480
 DB 1944 GATTAAGAGAAAGCCAGCATCCCAAGAGCTGCAAGACCGCAAGCTTCAAGACACCA 2003
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 DB 2004 AGCAAGCATTCATTCGGTCAAGAACCAAGCATGAGCAAGCGCCCAAGAGTCCCTTTTA 2063
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTyraHisThrSerPheLeuPheGly 520
 DB 2064 TCTCCACTGATCAAGTGGAGCGTGGAGGACAAATTAACACACCACTTCTTTCGGC 2123
 QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyraHis 540
 DB 2124 CTTCACCAAGCAGAGACCTCAAGAGTGTGCTGGCTTAAGGCTGAC 2183
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyraPheAla 560
 DB 2184 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTAACAGAGCTGATTTTGCC 2243
 QY 561 ThrGlnSerSerHisPheTyraSerAlaSerAlaIleTyraGlySerAlaSerTyraSer 580
 DB 2244 ACAGAGTCTTCACTTACTGCTGCTGACGATCTTAAGAGGAGCTCCAGTTACTCT 2303
 QY 581 AlaTyraSerCysSerGlnLeuProThrCysGlyAspGlnValTyraSerValArgArgArg 600
 DB 2304 GCTTACAGTGGACGACGCTGCCACTTGGAGAGCAAGCTATTTCTGTGGCAGGCGG 2363
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyraIleGluGluSerProPheGlu 620
 DB 2364 CAGAGCCAGAGTACAGAGCTACTGCGGCGAGCTGCGATGAGAGAGCCCTTTGAA 2423
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
 DB 2424 AAGCAGTTTAAACGCAAGAGCTGCAAAATGAAATTTGAGAGACATCATGTCAAGAAC 2483
 QY 641 ArgSerArgGluGluGluGlyValValGlySerGlnSerSerPheSerGlyLysMetGlu 660
 DB 2484 AGGTCAAGGAGAGCTGGGAGAAAGTGGGCACTGATCTTTCGGGAGCATGAGAA 2543
 QY 661 IleIleGluValSer 665

Qy 461 AsplAglyGluAlaSerIleProlyLeuGlnThrAlaArgProSerAspSerGln 480
 Db 1969 GATAGAGAGAAAGCCAGCATCCCAAGAACTGCAGACCCCGAGCTTCAGACAGCAG 2028
 Qy 481 SerIleArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 Db 2029 AGCAAGAGATTGATCGGTGAGAACAGAGAGTGGACCCGCCAGAGGTCCCTTTTA 2088
 Qy 501 SerProLeuHisArgSerGlySerValGluAspAsnThrIleThrSerPheLeuPheGly 520
 Db 2089 TCCTCCATCGATCGAAGTGGAGCGTGGAGAACATTTACACACACAGCTTCCTTTCCGC 2148
 Qy 521 LeuSerThrSerGlnGlnHisLeuThrLySerSerAlaGlyLeuGlyLeuLySerGlyTyrHis 540
 Db 2149 CTTTCCACCGACGACGACCTCAGAAATGCTGCTGCTGAGCTTAAGGCTGGCAC 2208
 Qy 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
 Db 2209 TCGGATATCTGGCCCCCGACAGCTTACCTCCCTTCCCTGACAGAGCTGTATTTTGC 2268
 Qy 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
 Db 2269 ACAGAGCTCCACACATCTTACTGCTGCTCAGCCATCTACGAGGCACTGCGAGTTACTCT 2328
 Qy 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
 Db 2329 GCTTACAGCTGACGACGCTGCCCACTTGGCGAGACCAATCTATTTCTGCGCAGCGG 2388
 Qy 601 GlnLySerProSerAspArgAlaAspSerArgArgSerTyrHisGluGluSerProPheGlu 620
 Db 2389 CAGAGCGCAAGTAGACAGAGCTGACTGCGCGGAGCTGGCATGAAGAGAGAGAGAGAGAG 2448
 Qy 621 LySerGlnPheValArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
 Db 2449 AACGAGTTTAAAGCGAGAACTGCCAATGAAATTTGGAGAGAGCATCATGTCAGAGAAC 2508
 Qy 641 ArgSerArgGluGluLeuGlyLyValGlySerGlnSerSerPheSerGlySerMetGlu 660
 Db 2509 AGGTACAGGAGAAAGCTGGGGAAGTGGGAGTCACTGCTTTTGGGACAGATGGA 2568
 Qy 661 IleIleGluValSer 665
 Db 2569 ATCATTTAGAGTCTCC 2583
 RESULT 10
 AF506796 3566 bp mRNA linear PRI 30-OCT-2003
 LOCUS Homo sapiens dual specificity phosphatase 16 (DUSP16) mRNA,
 DEFINITION complete cds.
 ACCESSION AF506796
 VERSION AF506796.1 GI:25573087
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3566)
 AUTHORS Hoonraet, I., Marynen, P., Goris, J., Sciort, R. and Baens, M.
 TITLE MAPK phosphatase DUSP16/MKP-7, a candidate tumor suppressor for
 chromosome region 12p12-13, reduces BCR-ABL-induced transformation
 JOURNAL Oncogene 22 (49), 7728-7736 (2003)
 PUBMED 14586399
 REFERENCES 2 (bases 1 to 3566)
 AUTHORS Hoonraet, I., Marynen, P. and Baens, M.
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-2002) Department for Human Genetics-Flanders
 Interuniversity Institute for Biotechnology (VIB), Katholieke
 Universiteit Leuven, Herestraat 49, Leuven B-3000, Belgium
 FEATURES
 source
 1. 3566
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

gene
 CDS
 /chromosome="12"
 /map="12p13-p12"
 /tissue_type="bone marrow"
 1. 3566
 /gene="DUSP16"
 /note="synonym: MKP7"
 633. 2630
 /gene="DUSP16"
 /EC_number="3.1.3.48"
 /EC_number="3.1.3.16"
 /codon_start=1
 /product="dual specificity phosphatase 16"
 /protein_id="AA075120.1"
 /db_xref="GI:25573088"
 /translation="MAHEMIGQIVTERIVALLSESTREKVLIDNSRPFVENVNTHILE
 AININCSKMKRRLOQDKVLTLEIQLSHKAKHDIDCSQKVYVTDSSQDVASLSSDC
 FLTVLLGLKESFNSVHLAAGPAERFSPGLCEKSTLVPCLISQPLPANIAPT
 RILPMLYLCCORDVNLKELMOQNGIGVYVNASNTCEKPDPIEPHSLRYPVDSFCEK
 ILPMLDSVDFLEKAKASNGCYLVHCLAGISRSATIAIAYIMKRMDSIDEAYRKYE
 KRPTLSNPNFPGOLIDYERKKIKNOTGASGPKSKLHLKRNKPEVPVASEGQKSE
 TPLSPCADSATSBAGQRPVHPASVPSPVQPSLLEDSPLVQALSGHLNSPRLED
 SNLKRFSFLDIKSVISASMAASLHGFSSSEDALRYTPSTLIDGTNKLQGPSVQV
 LSEQPTSPDKEBASIPKQLQYARSDSQSLRHSVRSSGTAGRSLLSLPLHRSQS
 VEDNHTSFLFGLSTSOQHILTKSAGLGLKQWHSDLIAPQTSPLTSLSWYFAESSHF
 YASATYGGASAYSAYSQSLPTCGPOVYSVRROKPSRADRSRMHESPEKQFK
 RSCQMEPESIMSERNSRELKXVGSOSFSGSMELIIVS"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,38e-204 Length: 3566
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: 9 Gaps: 0
 US-10-029-345A-109 (1-665) x AF506796 (1-3566)
 Qy 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 Db 633 ATGGCCCATGAGATGATGAACTCAAAATGTTACTAGAAAGTTGGTGGCTGCTGCGAA 692
 Qy 21 SerGlyThrGluLyValLeuLeuIleAspSerArgProPheValGluTyrThrAsnThrSer 40
 Db 693 AGTGAACCGAANAAGTCTGCTAATGATGATGACCGGCCCATTTGTGGAATTAACATACATCC 752
 Qy 41 HisIleLeuGluAlaIleAsnIleAsnCySerLyLeuMetLyArgArgLeuGlnGln 60
 Db 753 CACATTTTGAAGCCATTAAATCACTGCTCCAACTTATGAAGGAAAGTTGCAACAG 812
 Qy 61 AsplLyValLeuIleThrGluLeuIleGlnHisSerAlaLyHisLyValAspIleAsp 80
 Db 813 GACAAAGTGTAAATACAGAGCTCAACGACATTCAGCAAAACATTAAGGTGACATTGAT 872
 Qy 81 CysSerGlnLyValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 Db 873 TGCAGTCAAGAGGTGTAGTTTACATCAAGAGCTCCCAAGATGTCTCTCTCTTCA 932
 Qy 101 AspCysPheLeuThrValLeuLeuGlyLyLeuGluLySerPheAsnSerValHisLeu 120
 Db 933 GACTGTTTCTCACTGCTTCTGAGTAACAGGAAGAGCTTCAACTCTGTTCACCTG 992
 Qy 121 LeuAlaGlyLyPheAlaGluPheSerArgCysPheProGlyLyLeuGlyGluLyLySer 140
 Db 993 CTTCAGAGGGGGTGTGCTAGTCTCTCGTTTCCCTGGGCTGTGAAGGAAATCC 1052
 Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 Db 1053 ACTTAGTCCCTCACTGCACTTCTGAGCTTGTCTTACTGTTCCCAACTTGGGCCAAC 1112
 Qy 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLyGluLeuIle 180
 Db 1113 CGAATCTTCCCAATCTTATCTTGGCTGCGAGCGAGATGTCCTCAACAGAGGCTCGAG 1172


```

QY      181  GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProIysProAspPhe 200
Db      1173 CAGAGAAATGGATTTGGTTATGGTTAAATGGCAGAAATACCTGTCCAAAGGCTGACCTT 1232
QY      201  IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLeuIleu 220
Db      1233 ATCCCGAGCTCATTTCTGGGTGGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 1292
QY      221  ProTrpLeuAspLysSerValAspPheIleGlyValAlaValAsnGlyCysVal 240
Db      1293 CGGTGGTTGGCAAAATCAGTACGATTTTCATGAGAAAGCAAAAGGCTCCAAATGATGTTG 1352
QY      241  LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db      1353 CTAGTGCACTGTTTGTGCTGGAGATCTCCGCTCCGCAACATCGCTATGCGCTACATCATG 1412
QY      261  LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIlysgIuLysArgProThr 280
Db      1413 AAGAGATGGACATGCTTTGATGACAACTTACAGATTTGTGAAGAAAGAAAGCCCTACT 1472
QY      281  IleSerProAsnPheAsnPheLeuGlyValLeuLeuAspLysArgIuLysIleLysAsn 300
Db      1473 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGACACTATGAGAAAGATTTAAGAAC 1532
QY      301  GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGluLysProAsn 320
Db      1533 CAGACTGAGGATCAGAGGCCCAAGAGCAAACTCAAGCTGCTGCACTGGAGAAAGCCAAAT 1592
QY      321  GluProValProAlaValSerGluGlyValLysSerGluThrProLeuSerProPro 340
Db      1593 GAACTGTCTCCCTGCTGTCTCAGAGGGTGGACAAAGAAAGAGAGCGCTCTCAGTCCACC 1652
QY      341  CysAlaAspSerAlaThrSerGluAlaAlaGlyValArgProValHisProAlaSerVal 360
Db      1653 TGTCGCCACTGCTGACTCTCAGAGGACAGAGCAAAAGCCGCTGATCCGCCACGCTTG 1712
QY      361  ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
Db      1713 CCCAGCGTGGCCAGCGGTGAGCCGCTGTGATGAGAGAGCGCCGCTGTCAGAGCGCTC 1772
QY      381  SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
Db      1773 AGTGGGCTGCACTGTCCGCAAGCGGTGAGAAAGCACAATAGCTCAACGCTTCTCTTC 1832
QY      401  SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
Db      1833 TCTCTGATATCAAAATCAGTTTCATATTCAGCCACATGSCAGATGCTCTTACATGCGCTTC 1892
QY      421  SerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
Db      1893 TCTCTATCAAGAAATGCTTTGGAAATACACAAACCTTCACTACTCGAATGGAGACCAAC 1952
QY      441  LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
Db      1953 AAGCTATGCAAGTTCTCCCTGTTCCAGAACTATCGAGACAGATCCCGCAACACACTCTC 2012
QY      461  AspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
Db      2013 GATTAAGAGAGAAAGCAGCATCCCAAGAAAGCTGACAGCCGACGCTTCAAGACACCAAG 2072
QY      481  SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db      2073 AGCAAGGATTTGCTTGGCTCAGAACCAAGCAAGAGTGGCACCCGCAAGAGGTCTCTTTTA 2132
QY      501  SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
Db      2133 TCTCTACATGATCGAAGTGGAGAGCGTGGAGACAAATTAACAACACAGCTTCTTTTGAGC 2192
QY      521  LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysGlyTyrHis 540
Db      2193 CTTTCCACCAAGCAAGCACTCAAGAAAGTGTGCTGGCTGAGGCTTAAAGGGCTGAGCAC 2252

```

```

QY      541  SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
Db      2253 TCGGATATCTTGGTGGCCCAAGACCTTACCTTCCCTGACACAGACACTGATTTTGCC 2312
QY      561  ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
Db      2313 ACAGAGCTCTCACTTCTTACTTCTGCTTCAAGCACTTACGAGAGAGAGTGCAGCTTACTCT 2372
QY      581  AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
Db      2373 GCTTACAGCTTCACACCTGCTGCCCTCCACTTGGAGAGCAAACTATTTCTGTGCCAGGCG 2432
QY      601  GlnLysProSerAspArgAlaAspSerArgArgSerThrHisGluGlnSerProPheGlu 620
Db      2433 CAGAAAGCAAGTGCACAGAGCTGACTCGCGGAGAGCTGGAGATGACAGAGAGCCCTTTGAA 2492
QY      621  LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlySerIleMetSerGluAsn 640
Db      2493 AAGCAGTTTAAAGCAGAAAGCTCCAAATGGAATTTGGAGAGACATCATGTCAAGAAC 2552
QY      641  ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db      2553 AGGTCAAGGAAAGAGCTGGGAGAAAGTGGCAGTCAAGTCTTTCGGGACAGCATGAA 2612
QY      661  IleIleGluValSer 665
Db      2613 ATCATTGAGGTCTCC 2627

```

RESULT 11
AX374994
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
Tang, Y. T., Elliott, V. S., Ramkumar, J., Yao, M. G., Burford, N.,
Wang, Y. E., Stewart, E. A., Gandhi, A. R., Patterson, C., Lee, E. A.,
Hafalla, A. J., Lu, D. A., Tribouley, C. M., Griffin, D. A., Baum, M. R.,
Yue, H., Warren, B. A., Nguyen, D. B. and Wallis, N. K.,
Protein phosphatases
Patent: WO 0210363-A 17 07-FEB-2002;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1..3766
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7480570CBI"

ORIGIN
Alignment Scores:
Pred. No.: 4,72e-204 Length: 3766
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: Gaps: 0

US-10-029-345A-109 (1-665) x AX374994 (1-3766)

```

QY      1  MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db      538 ATGGCCCAAGAGATGATTTGAAATCTCAAAATGTACTGAGAGTTGGTGGCTCTGCTGAA 597
QY      21  SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db      598 AGTGAAGCGAAGAAAGTGTCTGTAATGATGATGCCGACATTTGTGGAATCAATACATCC 657

```


Qy	41	HisIleuGluAlaIleAsnIleAsnCySerIleuMetLysArgIleuGln	60
Db	658	CACATTTGGAGCCATTATATCACTGCTCCAGATTATGAAGGAGTTGCAACG	717
Qy	61	AspLysValIleuIleThrGluIleuIleGlnHisSerAlaLysHisValAspIleAsp	80
Db	718	GACAAAGTGTATTATACAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTGAT	777
Qy	81	CysSerGlnLysValValIleYrAspGlnSerSerGlnAspValAlaSerLeuSerSer	100
Db	778	TGCAGTCAGAAAGTGTAGATTACGATCCAAAGCTCCCAAGATGTGCTCTCTCTTCA	837
Qy	101	AspCyPheLeuThrValIleuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu	120
Db	838	GACTGTTTCTCACTGACTCTGGGTAACCTGAGAAAGAGCTTCACTGTTCACCTG	897
Qy	121	LeuAlaGlyLysIlePheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyLysSer	140
Db	898	CTTGCAAGTGGTGTGCTGAGTCTCTCGTGTGTTCCCTGCTGTGAGGAAATCC	957
Qy	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	958	ACTCTAGTCCCTACCTGCACTTCTCAGCCCTTGCTTACTGTGCCAACATTGGGCCAAC	1017
Qy	161	ArgIleLeuProAsnLeuIleuGlyCysGlnArgAspValLeuAsnLysGluLeuIle	180
Db	1018	CGAATCTTCCCAATCTTATCTTGCTGCTGAGGAAATGCTCCCAACAAAGAGCTGATG	1077
Qy	181	GlnGlnAsnGlyIleGlyIleValIleuAsnAlaSerIleThrCysProLysProAspPhe	200
Db	1078	CACGAGAAATGGATGGTTATGTCTTAAATGCCAGCAATCCTGTCCAAAGCTGACTTT	1137
Qy	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu	220
Db	1138	ATCCCGAGCTCATTTCTGCGGTGCTGTGAATACAGACTTGTGTGAGAAATTTTG	1197
Qy	221	ProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal	240
Db	1198	CCGTGTGTGACAAATCAGTAGATTTCTATGAGAAAGCAAAAGCTCCCAATGATGTGTT	1257
Qy	241	LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet	260
Db	1258	CTAGTGCATCTTAACTGAGTGGATCTCCCGCTCCGCCACCATCGCTATCGCTACATCATG	1317
Qy	261	LysArgMetAspMetSerLeuAspGluAlaIleYrArgPheValLysGluLysArgProThr	280
Db	1318	AAGAGATGAGCATGTCTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAACCTACT	1377
Qy	281	IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIleGluLysLysIleLysAsn	300
Db	1378	ATATCTCCAAACTTCAATTTCTGGGCCAACTCTGAGCTATGAGAAAGAAATTAAAGAAC	1437
Qy	301	GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisIleLeuGlyLysProAsn	320
Db	1438	CAGACTGAGACATCAGGAGCCAAAGAGCAACTCAAGCTGCTGCACTCGAGAAAGCCAAAT	1497
Qy	321	GluProValProAlaValSerGluGlyGlyLysSerGlnThrProLeuSerProPro	340
Db	1498	GAACTGTCTCTGCTGTCTCAGAGGGTGAAGAGAAAGAGAGAGCCCTCAGTCCACCC	1557
Qy	341	CysAlaAspSerAlaThrSerGluAlaAlaGlyLysArgProValHisProAlaSerVal	360
Db	1558	TGTGCCGACTCTCTACTCAGAGGAGGAGAGCAAAAGGCCGCTGCACTCCGCGAGGTG	1617
Qy	361	ProSerValProSerValGlnProSerLeuGlnLysAspSerProLeuValGlnAlaLeu	380
Db	1618	CCCAAGCTGCCACGCGAGCCCTGCTGTAGAGACAGCCGCTGTGACAGCCCTC	1677
Qy	381	SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe	400
Db	1678	AGTGGCTGACCTGCTCCGAGAGAGGCTGGAAGACAGCAATAAGCTCAGAGCTTCTTC	1737
Qy	401	SerLeuAspIleLysSerValSerIleSerAlaSerMetAlaIleSerLeuHisGlyPhe	420
Db	1738	TCTCTGATATCAAAATCACTTTCATATTCAGCCAGATGGACATCTTACATGGCTTC	1797
Qy	421	SerSerSerGluAspAlaLeuGluIleYrLysProSerThrThrLeuAspGlyThrAsn	440
Db	1798	TCCTCATCAGAAAGATCTTGTGAAATCTACAAACCTTCCACTCTGTGTGGAGACCMAC	1857
Qy	441	LysLeuCyGlnPheSerProValGlnIleuSerGlnGlnThrProGluThrSerPro	460
Db	1858	AAGCTATGCAATCTTCTCCCTGTTCAGGAATATGAGAGAGCTCCGAAACCAAGTCT	1917
Qy	461	AspLysGluGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln	480
Db	1918	GATAGAGAGAACCCAGCATCCCAAGAGCTGCAACCCGAGCTTCAAGAGCCAG	1977
Qy	481	SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu	500
Db	1978	AGCAAGCATTTGATTCGGTCAAGAACCAAGAGAGTGGCACCCGCCAGAGTCCCTTTTA	2037
Qy	501	SerProLeuHisArgSerGlySerValGluAspAsnThrHisPheSerPheLeuPheGly	520
Db	2038	TCTCACTGATCGAAGTGGAGCTGTGAGAGCAATTACCAACCAAGCTTCTTTTCGGC	2097
Qy	521	LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyThrHis	540
Db	2098	CTTTCACCAAGCCAGCAGACCTTCAAGAGTCTGCTGGGCTTAAAGGCTGGCAC	2157
Qy	541	SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpIlePheAla	560
Db	2158	TGGATATCTTGGCCCCCAGACCTTACCCCTGCTGACAGAGCTGGTATTTTGGC	2217
Qy	561	ThrGluSerSerHisPheIleSerAlaSerAlaIleIleGlyGlySerAlaSerTrpSer	580
Db	2218	ACAGAGTCTTCACTTCACTTCACTTCTGCTCAGCATCTACGAGAGAGTGCAGTACTCT	2277
Qy	581	AlaIleSerCysSerGlnLeuProThrCysGlyAspGlnValIleYrSerValAlaArgArg	600
Db	2278	GCCTTACGCTGACGACGAGCTGCCCACTTGGAGAGCAAGTCTATTTCTGCGCAGCGG	2337
Qy	601	GlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGlnSerProPheGlu	620
Db	2338	CAGAACCAAGTACAGAGAGTGACTCGCGGCGAGCTGGCATGAAGAGAGCCCTTTGAA	2397
Qy	621	LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn	640
Db	2398	AACGAGTTTAAAGCAGAGCTGCAATGGAATTTGAGAGAGCATCATGTCAAGAGAAC	2457
Qy	641	ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu	660
Db	2458	AGGTCAAGGAGAGAGCTGGGGAAGTGGGAGTCACTTACGTTTTCGGGACATGGAA	2517
Qy	661	IleIleGluValSer 665	
Db	2518	ATCATTTAGAGTCTCC 2532	
RESULT 12			
BD171157			
LOCUS			
BD171157 4790 bp DNA linear PAT 17-JAN-2003			
DEFINITION			
Novel gene and protein encoded thereby.			
ACCESSION			
BD171157			
VERSION			
BD171157.1 GI:27876969			
KEYWORDS			
MO 02052005-A/13.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 4790)			
REFERENCE			
Obara, O., Nagase, T. and Nakajima, D.			
Novel gene and protein encoded thereby			
Patent: WO 02052005-A 13 04-JUL-2002;			
KAZUSU DNA RESEARCH INSTITUTE, OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE			
NAKAJIMA			
COMMENT			
OS Homo sapiens (human)			

QY	241	LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleTyrIleMet	26
Db	904	CTAGTSCACTGTTTAGTGGGATCTCCGGCTCCGACCAATGCGTATCGCTTACATCATG	963
QY	261	LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIysGlyLysArgProThr	280
Db	964	AAGAGGATGAGCATGCTTTAGATGAAGCTTACAGATTGTGGAAAGAAAAAGCCTACT	1020
QY	281	IleSerProAsnPhaAsnPhaLeuGlyGlnLeuLeuAspTyrGlyLysIleLysAsn	300
Db	1024	ATATCTCCAACTTCAATTTCTGGGCCACTCTGGACATATGAGAAAGATTAAAGAAC	1080
QY	301	GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlyLysProAsn	320
Db	1084	CAGACTGGAGCATCAGGGCCMAAAGCAAACTCAAGCTGTGCACCTGGAGAACCAAT	1140
QY	321	GlnProValProAlaValSerGlnGlyGlnLysSerGlyThrProLeuSerProPro	340
Db	1144	GAACTGTGCTCTGTCTCTCAGAGGGTGGACAAAAAGCAAGCGCCCTCAGTCCACCC	1200
QY	341	CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal	360
Db	1204	TGTGCCACTCTGCTATCTCAGAGGACAGACAGCAAAAGCCGTGATCCCCACGGTG	1260
QY	361	ProSerValProSerValGlnProSerLeuGluAspSerProLeuValGlnAlaLeu	380
Db	1264	CCGACGGTGCACAGCGTGCAGCCGTGCTGTAGAGAACAGCCCGCTGTATCAGCGCTC	1320
QY	381	SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe	400
Db	1324	AGTGGCGTGCACCTGTCCCAAGCAGGCTGGAAGACAGAAATAACTCAAGCTTCTTC	1380
QY	401	SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaIleSerLeuHisGlyPhe	420
Db	1384	TCTCTGGATATCAATCACTTTTCAATATTCAGCCACATGGACAGCATCTTTCATGCGCTTC	1440
QY	421	SerSerSerGluAspAlaLeuGlnTyrTyrLysProSerThrThrLeuAspGlyThrAsn	440
Db	1444	TCTCTATCAGAAAGTCTTTGGAAATCAACAACCTTCACTACTCTGGATGGAGCAAC	1500
QY	441	LysLeuCysGlnPheSerProValGlnIleLeuSerGlnThrProGlnThrSerPro	460
Db	1504	AAGCTATGCACTTCTCCCTGTTTGAACATATCGAGACAGCATCTCCGAAACAGAGCTT	1560
QY	461	AspLysGlnGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln	480
Db	1564	GATTAAGAGGAAGCAGCATCCCAAGAGCTGCACCCGAGGCTTCAAGACAGCAG	1620
QY	481	SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu	500
Db	1624	AGCAAGCATTCGATTCGATCAGAACACAGACAGTGGACCCGCCAGAGTCCCTTTTA	1680
QY	501	SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly	520
Db	1684	TCTCCACTGCATCGAAGTGGAGGTGGAGAGACATTAACACACCAAGCTTCTTTTCGGC	1740
QY	521	LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrHis	540
Db	1744	CTTTCCACCAAGCAGACCACTTCAGAAAGTCTGCGGCTGGGCTTAAAGGCTGGCAC	1800
QY	541	SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrPyrPheAla	560
Db	1804	TGGATATCTTGGCCCCCAGACTTACCCCTTCCCTACCAAGCAGCGGTATTTTGC	1860
QY	561	ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer	580
Db	1864	ACAGAGTCTCACAATTCTTACTGCTCTCAGCACTTACGAGGAGCAGTGCAGTTACTCT	1920
QY	581	AlaTyrSerCysSerGlnLeuProThrTyrGlyAspGlnValTyrSerValArgArgArg	600
Db	1924	GCTTACAGCTGACGACAGCTGCCACTTGGAGAACCAAGTCTATTTCTGTGCGAGCGGG	1980

QY 601 GlnlybProSerAspArgAlaAspSerArgSerTrpHisGluGluSerProPheGlu 620
DB 1984 CAGAGCCAACTGACAGAGCTGACTCCGGCGAGCTGGCATGAAGAGCCCTTTGAA 2043
QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGluSerIleMetSerGluAsn 640
DB 2044 AAGCAGTTTAAACGCAAGAGCTGCCAATGGAATTTGAGAGACATCATCTCAGAGAAC 2103
QY 641 ArgSerArgGluGluMetGluLysValGlySerGlnSerPheSerGlySerMetGlu 660
DB 2104 AGGTCAAGGAGAGAGCTGGGAGAGTGGCAGTCACTAGCTTTTGGGCGCAGCATGAA 2163
QY 661 IleGluValSer 665
DB 2164 ATCATTTAGGTCTCC 2178
RESULT 13
LOCUS BD183422 4790 bp DNA linear PAT 17-JUN-2003
DEFINITION Novel genes and proteins encoded by the genes.
ACCESSION BD183422
VERSION BD183422.1 GI:31875622
KEYWORDS JP 2002345492-A/135.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 4790)
AUTHORS Ohara,O., Nagase,T. and Nakajima,D.
TITLE Novel genes and proteins encoded by the genes
JOURNAL Patent: JP 2002345492-A 135 03-DEC-2002;
KAZUSA DNA RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
FN JP 2002345492-A/135
PD 03-DEC-2002
PI 26-FEB-2002 JP 2002049009
PR OSANU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/09,C07K14/47//A61K31/711,A61K38/00,A61K48/00,A61P25/00,
PC A61P25/14,
PC A61P25/18,A61P35/00,C12N15/00,A61K37/02
CC Novel genes and proteins encoded by the genes FH Key
Location/Qualifiers
FT CDS (184)..(2178).
FEATURES
Source Location/Qualifiers
1..4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6.61e-204 Length: 4790
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: 6 Gaps: 0
US-10-029-345a-109 (1-665) x BD183422 (1-4790)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 184 ATGGCCCATAGAGATGATGGAACTCAAAATGTTACTAGAGAGTGGGCTGCTGGAAA 243
QY 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyraAsnThrSer 40
DB 244 AGTGAACGGAAGAAAGTGTGCTAATTGATAGCCGCACTTGGAGATCAATACATCC 303
QY 41 HisIleLeuGluValAlaIleAsnIleAsnCysSerIleValMetLysValArgLeuGlnGln 60
DB 304 CACATTTTGAAGCATTAAATCATCTCCCAAGCTTATGAAGCGAAGTTGCAACAG 363
QY 61 AspLysValIleLeuThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80

DB 364 GACAAAGTGTATTATACAGAGCTCATCCAGCATTTACGCGAAACATAAGTTGACATTGAT 423
QY 81 CysSerGlnLysValValValIlyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 424 TGCAGTCAGAGAGGTGTAGTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 483
QY 101 AspCysPheLeuThrValLeuLeuGluLysLeuGluLysSerPheAsnSerValHisLeu 120
DB 484 GACTGTTTCTCACTGATCTTGGGTAACTGAGAGAGCTTCAACTGTTCACTG 543
QY 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuGluGlyLysSer 140
DB 544 CTTCGAGTGGGTGTTCTGCTGAGTCTCTCGTGTTCCTGCTCTGTAAGGAAATCC 603
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 604 ACTTAGTCCCTACCTGACATTTCTCAGCCCTTGCTTACCTGTCCAACATTTGGGCCAAC 663
QY 161 ArgIleLeuProAsnLeuTyrlLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
DB 664 CGAATTTCTCCCAATCTTATCTTGGCTGCACAGAGATGCTCTCAACAGAGCTGATG 723
QY 181 GlnGlnAsnGlyIleGlyTyrlValLeuAsnAlaSerTyrlThrCysProLysProAspPhe 200
DB 724 CACAGAGATGGATGTGTTATGTGTTAATGCCAGCAATCTGTCCAAAGCTGACTTT 783
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
DB 784 ATCCCGAGCTCATTTCCGCTGCTGCTGTAATGACAGCTTTGTGAGAAATTTTG 843
QY 221 ProThrLeuAspLysSerValAspPheIleGluValAlaLysSerAsnGlyCysVal 240
DB 844 CCGTGTGTGACAAATCAGATGATTTTATTGAGAAAGCAAAACCTCCAAATGATGTGT 903
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlMet 260
DB 904 CTAGTGCACGTGTTTACTGCTGAGACTCCGCTCCGCCACCATCTGCTCATCATG 963
QY 261 LysArgMetAspPheSerLeuAspGluAlaTyrlArgPheValLysGluLysArgProThr 280
DB 964 AAGAGATGAGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAGAACCTACT 1023
QY 281 IleSerProAsnPheAsnPheLeuGluGlnLeuLeuAspTyrlLysLysIleLysAsn 300
DB 1024 ATATCTCCCAACTTCAATTTCTGGGCCCACTCTGACATGAGAAAGATTAGAAC 1083
QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGluLysProAsn 320
DB 1084 CAGACTGAGACATCAAGGCCCAAGAGCAAACTCAAGCTGTGCACTTGAGAGCCCAAT 1143
QY 321 GluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProPro 340
DB 1144 GAACGTGCTCCTGCTGCTGCTCAGAGGGTGAAGAAAGCAGAGCCCTCACTCCACC 1203
QY 341 CysAlaAspSerAlaHisSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
DB 1204 TGTGCCACTGCTCTACCTCAGAGGACAGAGCAAAAGCCCTGTGATCCCGCACGCTG 1263
QY 361 ProSerValProSerValGlnProSerLeuLeuGluLysAspSerProLeuValGlnAlaLeu 380
DB 1264 CCGAGCTGCCCAAGCTGACGCTGCTGCTGTTAGAGACAGAGCCGCTGTGACAGGCTC 1323
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
DB 1324 AGTGGCTGCACTGTGCCAGACAGCTGGAAGACAGCAATAGCTCAAGCCTTCTTC 1383
QY 401 SerLeuAspIleLysSerValSerTyrlSerAlaSerMetAlaSerLeuHisGlyPhe 420
DB 1384 TCTCTGATATCAAAATCAGATTATATTCAGCCAGATGCGACATCTTACATGCTTC 1443
QY 421 SerSerSerGluAspAlaLeuGluTyrlTyrlLysProSerThrThrLeuAspGlyThrAsn 440

Db 1444 TCCTCATCAGAGAGTCTTGGAAATACACAAACCTTCACACTCTGATGGAGCAC 1503
QY 441 LysLeuCYGlnPhseSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
Db 1504 AGCATATGCAATCTTCTCCCTCTCTCAAGAACTATCGAGAGACTCCCGAAGACAGTCTT 1563
QY 461 AspGlyGlnGlnAlaSerLeuProGlyLeuGlnThrAlaArgProSerAspSerGln 480
Db 1564 GATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1623
QY 481 SerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArgSerLeu 500
Db 1624 AGCAGGCAATTCATTCCTGCTGAGAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1683
QY 501 SerProLeuHisArgSerGlySerValGlnAspAsnThrHisThrSerPheLeuPheGly 520
Db 1684 TCTCCACATCGATCGAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1743
QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyThrHis 540
Db 1744 CTTTCACACAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1803
QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrProPheAla 560
Db 1804 TCGGATATCTGGCCCCCAGACCTTACCCCTTCCCTGACACAGCTGGATATTTTGGC 1863
QY 561 ThrGlnSerSerHisPheThrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
Db 1864 ACAGAGTCTCTACACTTCTACTCTGCTCAGCCATCATCAGGAGGAGGAGGAGGAGGAGGAG 1923
QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerAlaArgArgArg 600
Db 1924 GCGTACAGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1983
QY 601 GlnLysProSerSerAspArgAlaAspSerArgArgSerTyrPheGlnGlnSerProPheGln 620
Db 1984 CAGAGGCAAGTACAGAGAGCTGACTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2043
QY 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlnGlnSerIleMetSerGlnAsn 640
Db 2044 AAGCAGATTAAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2103
QY 641 ArgSerArgGlnGlnLeuGlnGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
Db 2104 AGGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2163
QY 661 IleIleGlnValSer 665
Db 2164 ATCATTGAGGTCTCC 2178

RESULT 14
AB051487
LOCUS AB051487 4790 bp mRNA linear PRI 07-FEB-2001
DEFINITION Homo sapiens mRNA for KIAA1700 protein, partial cds.
ACCESSION AB051487
VERSION AB051487.1 GI:12697944
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Nagase,T., Kikuno,R., Hattori,A., Kondo,Y., Okumura,K. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XIX. The complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL DNA Res. 7 (6), 347-355 (2000)
MEDLINE 21082932
PUBMED 11214970
REFERENCE 2 (bases 1 to 4790)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,

Department of Human Gene Research: 1512-3, Yana, Kiseazuru, Chiba
292-0812, Japan (E-mail:cdna@fokkazusa.or.jp)
URL: <http://www.kazusa.or.jp/huge>, tel:81-438-52-3913,
Fax:81-438-52-3914

FEATURES
source
location/Qualifiers
1..4790
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fj15153"
/note="vector:pb1uescriptII SK plus"
1..4790
/gene="KIAA1700"
<109..2181
/gene="KIAA1700"
/note="Start codon is not identified."
/codon_start=1
/product="KIAA1700 protein"
/protein_id="BAB21791.1"
/db_xref="GI:12697945"
/translation="AFETHTITTSWQIEKGGRGLIYVMAHEMIGTOIVTERIVALL
ESGTEREVLIDSRPFVEVNTSHLEAININSGKMKRRLQODKVLTELIGSARKY
DIDSGKVVYDDSDQVSLSDPCLTVLLEKRSRSHILAGFAERSRCPGL
CEKSTLVPCTISQPLPYANIGPRILIPNLYLQCRDYNLELMQNGIGVPLVANS
ATTAIAYIKRMDMSDEAYRVKRPITSPENFLQQLDYKKIKINQTAGSPKS
KTLULHEKPNRPVPSVSGQKSETPSPCADATSEAGQRPVHPVSPVSVQ
PSLLEDSPLVQALSGHLASADRLSDNKLKSFSLDISVSAVSMASLHGFSSSD
ALFEYKPSITLDGNTKLCQFSPVQELSDOTPEPSXDEKESIPKLTOTARPDSQSR
LHSVRRSSSGTQKRSILSPHRSRGVENYHTSPFLRGSTSOHLTKSAGLIGMS
DILAPQSTPSLSTSSMYFATESHSHYVSAALYGGASASATSCQLPTCGQVSVSR
RQKSPDRARSRSWMBESPFERQFRKSCQMFESJWSENRSREELKGVSSQSSFSG
SMEIIEVVS"

ORIGIN
Alignment Scores:
Pred. No.: 6,61e-204 Length: 4790
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
Gaps: 0
US-10-029-345A-109 (1-665) x AB051487 (1-4790)

QY 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuGln 20
Db 184 ATGGCCCATGAGATGATTTGGAATCTCAAAATGTGTACTAGAGGTTGGTCTGCTGAGAA 243
QY 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnTyrAsnThrSer 40
Db 244 AGTGAACGGAAGAAAGGCTGCTTAATTGATTAGCGCGGCATTGTGGAAATCAATACATCC 303
QY 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerLeuLysLeuMetLysArgArgLeuGlnGln 60
Db 304 CACATTTGGAGGCATTAAATATCACTGCTCAAGCTTATGAAAGGAGTTGCAACAG 363
QY 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 364 GACAAAGGTTAATTAACAGAGCTCATCCAGCATTCAGCAACATAGAGTTACATGAT 423
QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 424 TGCAGTACGAGAGGTTGATGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 483
QY 101 AspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 484 GACTGTTTCTCATCTGATCTTCGGGTAACTGGAGAAAGCTTCAACTGTGTTACCTG 543
QY 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCYGlnGlyLysSer 140
Db 544 CTTGACAGGTTGGTTGCTGATGTTCTCTGTTGTTCCCTGGCTCTGTGAAGAAATCC 603

QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaIleuIleGlyProThr 160
 DB 604 ACTCTAGTCCCTTACCTGCAATTTCTCAGCCTTGCTTACCTTGGCAACATTTGGGCCAAC 663
 QY 161 ArgIleLeuProAnleuTyrlleuGlyCysGlnArgPheValLeuAnlysgIleuIle 180
 DB 664 CGAATCTTCCCAATCTTTATCTTGGCTGCAGAGAAATGCTCTCAACAGAGAGCTGATG 723
 QY 181 GlnGlnAnGlyIleGlyTyrlValLeuAnAlaSerTyrlCysProlyProAnPhe 200
 DB 724 CACAGAAATGGAAATGGTTATGGTTAAATGCAAGCAATACCTGTCMAAGCCTGACTTT 783
 QY 201 ILProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnTyrlleu 220
 DB 784 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGATACAGCTTTGTGGAATAATTTTG 843
 QY 221 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAnGlyCysVal 240
 DB 844 CCGTGGTTGGACAATCAGTAGATTTCATTGAGAAACAAAGCCTCCATGTGATGTT 903
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlIleMet 260
 DB 904 CTATGTGACATGTTAGCTGGGAATCTCCGCTCCGCAACATCGCTACATCATG 963
 QY 261 LysArgMetAspMetSerLeuAspGlnAlaTyrlArgPheValLysGlnLysArgProThr 280
 DB 964 AAGAGATGACATGTTTATGATGAACTTACAGATTTGTGAAAGAAAAGAACTTACT 1023
 QY 281 ILSerProAnPheAnPheAnPheLeuGlyGlnLeuAnPTrGlnLysLysIleLysAsn 300
 DB 1024 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGGACATATGAAAGAAATTAAGAAC 1083
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuAnlysgIleuGlyLysProAn 320
 DB 1084 CAGACTGAGCAATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTCGAGGAAGCCCAAT 1143
 QY 321 GlnProValProAlaValSerGlnGlyGlnLysSerGlnPheProLysSerProPro 340
 DB 1144 GAACCTGTCCCTGTCTCTCAGAGGGTGAACAGAAAGAGCAAGAGCCCTCAGTCAAGCC 1203
 QY 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 DB 1204 TGTGGCACTCTGTACTCTCAGAGGAGCAAGCAAAAGCCCGTGCATCCGCGCAGGTG 1263
 QY 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
 DB 1264 CCGAGCTGCGCCAGCGTGCAGCCGCTGTGTAGAGACAGCCGCTGTGACAGGCCCTC 1323
 QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAnlysgIleuLysArgSerPhe 400
 DB 1324 AGTGGGCTGCACTGTCCGAGAGAGGCTGAGAGACAGCAATAGCTCAAGCGTTCTTC 1383
 QY 401 SerLeuAspIleLeuSerValSerTyrlSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 DB 1384 TCTCTGGATTCAAATAGTTTCAATATTACGCAAGCATGAGATCTTACATGCTTC 1443
 QY 421 SerSerSerGlnAspAlaLeuGlnTyrlTyrlAspSerThrThrLeuAspGlyThrAsn 440
 DB 1444 TCTCTCATCAAGAAAGATGCTTGGAAATACTAACAACCTTCCACTGTGATGGAGCAAC 1503
 QY 441 LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnGlnPheProGlnThrSerPro 460
 DB 1504 AAGCTATGCAAGTTCCTCCCTGTTCAAGAACTATCGAGAGCACTCCGAAACCAAGTCT 1563
 QY 461 AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 DB 1564 GATTAAGAGAGAAAGCAACATCCCAAGAGCTGAGAGCCCGCAGGCTTACAGACAGCAG 1623
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 DB 1624 AGCAAGCGATGTGATTCGTGTGAGAAACAGAGCAGTGGCAGCCGCGCAGAGGTCTCTTTTA 1683
 QY 501 SerProLeuHisArgSerGlySerValGlnAspAnlyrlHisThrSerPheLeuPheGly 520

DB 1684 TCTCCACTCATCGAAGTGGAGGCGTGAAGCAATTAACACAGCACTTCCCTTCGCG 1743
 QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTrpHis 540
 DB 1744 CTTTCCACAGCAGCAGAGCACTCAAGAAATGCTGCGCTGGCCTTAAAGGCTGGCAC 1803
 QY 541 SerAspIleLeuAlaProGlnPheSerThrProSerLeuThrSerSerTrpTyrlPheAla 560
 DB 1804 TCGGATATCTGGCCCCCAAGACTTACCTTACCCTTCCCTGACAGCACTGGATTTTGGC 1863
 QY 561 ThrGlnSerSerHisPheTyrlSerAlaSerAlaIleTyrlGlySerAlaSerTyrlSer 580
 DB 1864 ACAGACTCTCACTTACTTACTGCTGCTCAGCATCTTAAGAGAGGAGTGCACATTA 1923
 QY 581 AlaTyrlSerCysSerGlnLeuProThrCysGlyAspGlnValTyrlSerValAlaGlyArg 600
 DB 1924 GCCTACAGCTGACGCGAGCTGCCCTTCCGAGAACAAAGTCTATTTCTGTGCCAGGCG 1983
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGlnGlnLysSerProPheGlu 620
 DB 1984 CAGAACCCAAAGTACAGAGCTGACTGCGCGAGCTGGCATGAAAGAACCCCTTTGAA 2043
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
 DB 2044 AAGCACTTTAAACGAGAAAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCAAGAAC 2103
 QY 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 DB 2104 AGGTCAAGGGAAGAGCTGGGGAAGAGTGGGCAAGTCACTTTCGGGCAAGCATGAA 2163
 QY 661 IleIleGlnValSer 665
 DB 2164 ATCATTTAGAGTCTCC 2178
 RESULT 15
 AX713989 2102 bp DNA linear PAT 15-Apr-2003
 LOCUS AX713989
 DEFINITION Sequence 673 from Patent EP1293569.
 ACCESSION AX713989
 VERSION AX713989.1 GI:29888917
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S., Yamamoto,J.I., Igono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and Masuno,Y.
 Full-length cDNAs
 Patent: EP 1293569-A 673 19-MAR-2003;
 Helix Research Institute (JP) ; Research Association for Biotechnology (JP)
 FEATURES
 source location/Qualifiers
 1..2102
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,74e-204 Length: 2102
 Score: 3399.00 Matches: 662
 Percent Similarity: 99.70% Conservative: 1
 Best Local Similarity: 99.55% Mismatches: 2
 Query Match: 99.44% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-345A-109 (1-665) x AX713989 (1-2102)
 QY 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: June 21, 2004, 21:18:06 ; Search time 819.731 Seconds

(without alignments)
3716.370 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 3418
Sequence: 1 MAHEMIGTQIVTRVALLE.....LKGVSQSSPSSGSMETIEVS 665

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 3017426 segs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seg length: 0
Maximum DB seg length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10029345/runcat 21062004.16416 29071/app_query.fasta_1.1294
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -INIT5=bits -START=1 -END=370 -MATRIX=blosum62
-TRANS=human40.cst -LIST=45 -DOCAIGN=200 -THR.SCORE=pcr -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADING=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10029345 @CGN_1_1_33 @runcat 21062004.16416_29071
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBBLCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPEXT=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	3406	99.6	1998	6	US-09-816-494-3	Sequence 3, Appl 1
2	3406	99.6	1998	5	US-10-377-072-27	Sequence 27, Appl 1
3	3406	99.6	2732	13	US-10-168-506-2	Sequence 2, Appl 1
4	3406	99.6	3059	17	US-10-257-026-1	Sequence 1, Appl 1
5	3406	99.6	3496	9	US-09-964-277-1	Sequence 1, Appl 1
6	3406	99.6	3544	9	US-09-816-494-1	Sequence 1, Appl 1
7	3406	99.6	3544	16	US-10-377-072-25	Sequence 25, Appl 1
8	3406	99.6	3625	13	US-10-425-114-26234	Sequence 26234, A
9	3406	99.6	3766	13	US-10-343-357-17	Sequence 17, Appl 1
10	3406	99.6	4790	17	US-10-648-559-115	Sequence 115, Appl 1
11	3399	99.4	2102	16	US-10-096-749-673	Sequence 673, Appl 1
12	3398	99.4	2966	13	US-10-294-115-520	Sequence 520, Appl 1
13	3379.5	98.9	2071	13	US-10-072-012-257	Sequence 257, Appl 1
14	3358.5	98.3	2200	13	US-10-072-012-255	Sequence 255, Appl 1
15	3048.5	89.2	3332	9	US-09-964-277-20	Sequence 20, Appl 1
16	1643.5	48.1	1916	16	US-10-108-260A-2429	Sequence 2429, Ap
17	1302	38.1	2453	14	US-10-005-858-1	Sequence 1, Appl 1
18	1262	36.9	2476	13	US-10-220-120-28	Sequence 28, Appl 1
19	1075.5	31.5	2039	13	US-10-072-012-265	Sequence 265, Appl 1
20	480	14.0	2109	16	US-10-305-720-1135	Sequence 1135, Ap
21	480	14.0	2109	17	US-10-641-643-946	Sequence 946, App
22	480	14.0	2390	13	US-10-342-887-698	Sequence 698, App
23	480	14.0	2390	13	US-10-172-118-658	Sequence 658, App
24	480	14.0	2390	15	US-10-175-523-183	Sequence 183, App
25	480	14.0	2390	16	US-10-199-221-4	Sequence 4, Appl 1
26	480	14.0	2380	17	US-10-717-597-4	Sequence 4, Appl 1
27	480	14.0	2792	10	US-09-971-392-69	Sequence 69, Appl 1
28	479	14.0	2794	14	US-10-044-090-330	Sequence 330, App
29	477	14.0	2649	16	US-10-133-937-48	Sequence 48, Appl 1
30	477	14.0	2649	16	US-10-159-563-48	Sequence 48, Appl 1
31	471	13.8	2650	15	US-10-184-832-1	Sequence 1, Appl 1
32	469	13.7	1337	9	US-09-964-899-46	Sequence 46, Appl 1
33	469	13.7	1449	12	US-10-052-482-192	Sequence 192, App
34	469	13.7	1830	15	US-10-346-356-1	Sequence 1, Appl 1
35	469	13.7	1909	12	US-10-052-482-191	Sequence 191, App
36	469	13.7	1999	13	US-10-058-270A-129	Sequence 129, App
37	468	13.7	1452	12	US-10-052-482-189	Sequence 189, App
38	468	13.7	1977	12	US-10-052-482-188	Sequence 188, App
39	466.5	13.6	1143	15	US-10-184-832-3	Sequence 3, Appl 1
40	466.5	13.6	2104	12	US-10-152-319A-5041	Sequence 2041, Ap
41	466.5	13.6	2104	16	US-10-388-934-654	Sequence 654, App
42	466.5	13.6	2104	16	US-10-191-803-270	Sequence 270, App
43	461	13.5	3286	10	US-09-971-392-9	Sequence 9, Appl 1
44	453	13.3	1238	9	US-09-736-457-803	Sequence 803, App
45	453	13.3	1238	9	US-09-902-941-803	Sequence 803, App

ALIGNMENTS

RESULT 1
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: PHOSPHATASE MOLECULES AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-816-494-3

Alignment Scores: 0
Pred. No.: 1998
Length: 1998

Score: 3406.00
 Percent Similarity: 99.85%
 Best Local Similarity: 99.70%
 Query Match: 99.65%
 DB: 9

Matches: 663
 Conservative: 1
 Mismatches: 1
 Indels: 0
 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-816-494-3 (1-1998)

QY 1 MetalAH1SGIleuMetIleGIYThrcGlnIleValThrcIuArgLeuValAlaleuLeuGlu 20
 Db 1 ATGGCCCATGAGATGATTGAACTCAAAATGGTTACTGAGAGGTGGTGGCTCTGGCGAA 60
 QY 21 SerGIYThrcGIuLySValIleuLeuIleAspSerArgProPheValGIuLyTrAsnThrSer 40
 Db 61 AGTGAACGGAAAAAGTGTGCTAATTGATAGCCGGCCATTGTGGAAATACAAATACATCC 120
 QY 41 HisIleuGluIuAlaIleuAlaIleuLeuCySerLySLeuMetLySArgArgLeuGln 60
 Db 121 CACATTTTGGAAACCATTAATATCAACTGCTCCAAAGCTTATGAAAGCAAGGTTGCACACG 180
 QY 61 AspLySValIleuIleThrcIuLeuIleGlnHisSerAlaLySHisLySValAspIleAsp 80
 Db 181 GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
 QY 81 CySerGlnLySValIleuValIleuValIleuValIleuValIleuValIleuValIleu 300
 Db 241 TGCAGTCAGAGGTTGATGATTACAGATCAAAAGCTCCCAAGATGTGGCTCTCTCTCA 300
 QY 101 AspPhePheLeuThrcValIleuLeuGlyLySLeuGluLySLeuPheAsnSerValHisLeu 120
 Db 301 GACTGTTTCTCAGTACTCTCTGGGTAACTGGAGAAAGCTTCAACTGTGTCACTG 360
 QY 121 LeuAlaGIYGIYPhelAGluPheSerArgCySerPheProGlyLeuCySGluGIYLySer 140
 Db 361 CTTCGAGAGGGGTTTGCAGATTCTCTCGTGTGTCTCTGCTGCTGTGAGAAATCC 420
 QY 141 ThrLeuValProThrCySIIeSerGlnProCySLeuProValAlaAsnIleGIYProThr 160
 Db 421 ACTCTAGTCCCTACCTCATTTCTCAGCTTGCCTTGCCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 161 ArgIleuLeuProAsnLeuIleuGIYLeuGIYCySLeuArgAspValIleuAsnLySLeuLeu 180
 Db 481 CGAATTTCTCCCAATCTTTTATCTTGGCTGCGACGAGATGCTCTCAACAGAGGCTGATG 540
 QY 181 GlnIleuAsnGIYIleGIYTrValIleuAsnAlaSerTrThrCySProLySProAspPhe 200
 Db 541 CAGCAGATGGATGGTATGTTAATGCCAGCAATACCTGTCCAAGCTGACTTT 600
 QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCySGluLySLeu 220
 Db 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 QY 221 ProTrpLeuAspLySLeuValAspPheIleGIYLySAlaLySAlaSerAsnGIYCySVal 240
 Db 661 CCGTGTGTGCAAAATCAGTACATTTCAATGAGAAAGCAAAAGCTTCATGATGGTGT 720
 QY 241 LeuValHisCySLeuAlaGIYIleSerArgSerAlaThrIleAlaIleAlaIleAlaIle 260
 Db 721 CTAGTGAAGTGTGTTAGCTGGGATCTCCGCTCCGCAACATCCCTATCCGCTCATCATG 780
 QY 261 LysArgMetAspMetSerLeuAspGluAlaIleuValIleuValIleuValIleuValIleu 840
 Db 781 AAGAGATGACATGTCTTATGATGAGAGCTTACAGATTGTGAAAGAAAAAGAAAGCTCAT 840
 QY 281 IleSerProAsnPheAsnPheLeuGIYGIuLeuLeuAspTrGIYLySValIleLySAsn 300
 Db 841 ATATCTCCAAACTTCATTTTCTGGCGCAACTCTGACATACAGAAAGAAAGATTAAAGAC 900
 QY 301 GlnThrcGIYAlaSerGIYProLySLeuLySLeuLySLeuLySLeuLySLeuLySLeu 320
 Db 901 CAGACTGAGATCAAGGCGCAAGAGCAAACTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 321 GluProValProAlaValSerGIYGIYGIYGIYGIYGIYGIYGIYGIYGIYGIYGIYGIYGIY 340

Db 961 GAACCTGTCCCTGCTCTCTCAAGAGGTGACGAAAAACGAGAGCCCTCTACGTCAACCC 1020
 QY 341 CysAlaAspSerAlaThrSerGIuAlaAlaGIYGIuArgProValHisProAlaSerVal 360
 Db 1021 TGTGCGACTCTGTACTTACCTCAGAGGCGACAGCAAAAGGCCCTGTGATCCCGCAGGTG 1080
 QY 361 ProSerValProSerValGIuProSerLeuLeuIleuAspSerProLeuValGlnAlaLeu 380
 Db 1081 CCAGCGGTGCCAGGCTGACCGCTGCTGTATGAGACAGCCCGCTGTATCAGCGCTTC 1140
 QY 381 SerGIYLeuHisIleuSerAlaAspArgLeuGIuAspSerAsnLySLeuLySArgSerPhe 400
 Db 1141 AGTGGGCTGCACTGCTCCGACAGACAGGCTGGAAGACAGCAATAGCTCAAGCGTTCCTTC 1200
 QY 401 SerLeuAspIleLySLeuValIleuTrpSerAlaSerMetAlaAlaSerLeuHisGIYPhel 420
 Db 1201 TCTCTGATATCAAAATAGTTTCAATATTCAGCCAGATGCGACATCTTACATGGCTTC 1260
 QY 421 SerSerSerGIuAspAlaLeuGIYTrpLySProSerThrThrLeuAspGIYThrcAsn 440
 Db 1261 TCTCATCAAGAAAGATGCTTGGAAATCTCAAAACCTTCCATCACTCTGATGGACACAC 1320
 QY 441 LysLeuCySGlnPheSerProValGIuGIuLeuSerGIuGIuTrpProGluThrSerPro 460
 Db 1321 AAGCTATGCCAGTTCCTCCCTGTTCAAGAACTATCGAGAGCACTCCGAAACCAAGTCT 1380
 QY 461 AspLySGIuGIuAlaSerIleuProLySLeuGIuTrpAlaArgProSerAspSerGln 480
 Db 1381 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGAGACCGCCAGGCTTCAGACAGCGAC 1440
 QY 481 SerLySArgLeuHisSerValArgThrSerSerSerGIYThrcAlaGlnAspSerLeuLeu 500
 Db 1441 AGCAACGATTCATTCGCTCAGAACCAAGCAAGCTGACACCGCCAGAGTCCCTTTTA 1500
 QY 501 SerProLeuHisArgSerGIYSerValGIuAspAsnTrpHisThrSerPheLeuPheGIY 520
 Db 1501 TCTCACTGCAATCAAGTGGAGGCTGAGAGACATTAACACACGCTTCCTTTTCCGAC 1560
 QY 521 LeuSerThrSerGIuGlnHisIleuThrLySLeuAlaGIYLeuGIYLeuLySGIYTrpHis 540
 Db 1561 CTTCCTCAACACAGCAGCACTCTCAAGAACTCTGCTGGGCTGAGGCTTGAAGGCTGGCAC 1620
 QY 541 SerAspIleuAlaProGlnThrThrSerThrProSerLeuThrSerSerTrpLyPheAla 560
 Db 1621 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGCAGCTGATTTTCC 1680
 QY 561 ThrGluSerSerHisPheTrpSerAlaSerAlaIleuGIYGIYLySLeuAlaSerTrpSer 580
 Db 1681 ACAGAGTCTCAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 QY 581 AlaTrpSerCySLeuGIuLeuProThrCySGlyAspGlnValIleuSerValArgArg 600
 Db 1741 GCTTACAGCTGCAAGCGAGCTGCTCACTTGGGAGACCAAGCTATTTCTGCTGCGCAGCGG 1800
 QY 601 GlnLySProSerAspArgAlaAspSerArgArgSerTrpHisGIYGIuLeuProPheGln 620
 Db 1801 CAGAACCAAGGACAGAGCTGACCTGCGGCGAGCTGGCAATGAAGAGACCCCTTTGAA 1860
 QY 621 LysGlnPheLySArgArgSerCySGlnMetGluPheGIYGIuLeuIleuSerGIuAsn 640
 Db 1861 AAGCATTTTAAACGAGAAAGCTGCGCAATGGAATTTGAGAGAGCATCATGTCAGGAAC 1920
 QY 641 ArgSerArgGIuGIuLeuGIYLySValGIYSerGIuSerSerPheSerGIYSerMetGlu 660
 Db 1921 AGGTACCGGAAAGAGCTGGGAAAGTGGGACAGCTCACTTCTTCTTCTGCGCAGCATGAA 1980
 QY 661 IleIleGIuValSer 665
 Db 1981 ATCATATGAGTCTCC 1995

RESULT 2
 US-10-377-072-27

Sequence 27, Application US/10377072
Publication No. US20040009501A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals Inc.
APPLICANT: Curtis, Rory A. J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Gluckmann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Miyoung
APPLICANT: Teal, Fong-Ying
TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: MP03-0180NMIM
CURRENT FILING DATE: US/10/377,072
CURRENT FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/895,860
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215,370
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/723,806
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/187,455
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/843,297
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,801
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: US 09/861,801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205,508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/816,494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/815,419
PRIOR FILING DATE: 2001-03-22
Remainder Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 1998
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1998)
US-10-377-072-27
Alignment Scores:
Pred. No.: 0 Length: 1998
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
Gaps: 0
US-10-029-345a-109 (1-665) x US-10-377-072-27 (1-1998)
QY 1 MetAlaHisGluWetIleGIYThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 1 ATGGCCCATAGATGATTGGAACCAAAATGTTACTGAGAGTTGGTGGCTCTGCTGGAA 60
QY 21 SerGIYThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 61 AGTGGACGGAAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 41 HisIleLeuGluAlaIleAsnIleAsnCySerIleuLeuLeuLeuLeuLeuLeuLeu 60
DB 121 CACATTTTGGAGGCAATTAATATCACTGCCCAAGCTTATGAAGCGAAGTTGCAACAG 180
QY 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80

DB 181 GACAAAGCTGTAATTCAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 240
QY 81 CySerGlnIleValIleValIleValIleValIleValIleValIleValIleValIle 100
DB 241 TGCAGTCAGAGGTTAGTTTACGATCAAGCTCCCAAGATTTGCTCTCTCTTCA 300
QY 101 AspCyPheLeuThrValLeuLeuGluIleValIleValIleValIleValIleValIle 120
DB 301 GACTGTTTCTCACTACTCTGAGTTTAACTGGAAGAGCTTCAACTGTTCACTG 360
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCyPheProGlyLeuCyGluGluIleVal 140
DB 361 CTTCGAGGAGGAGTTCTGATGTTCTGCTGTTTCCCTGCTGCTGCTGCTGCTGCTGCT 420
QY 141 ThrLeuValProThrCySilesGlnProCySleuProValAlaAsnIleGlyProThr 160
DB 421 ACTCTAGTCCCTACCTGACATTTCTCAGCCCTGCTTACCTGTTCCAAATGGGCCAAC 480
QY 161 ArgIleLeuProAsnLeuIleValIleValIleValIleValIleValIleValIleVal 180
DB 481 CGAATCTTCCCAATCTTATCTTGGCTGCGACGAGATGCTCTCAACAGAGCTGATG 540
QY 181 GlnGlnAsnGlyIleGlyIleValIleAsnAlaSerIleThrCyAspProIleProAspPhe 200
DB 541 CAGCAGAAATGGATTTGTTATGCTTAAATGCCAGCAATACCTGTCMAAGCTGACTTT 600
QY 201 IleProGluSerIlePheLeuArgValProValAsnAspSerPheCyGluIleVal 220
DB 601 ATCCCGAGCTCATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 221 ProThrLeuAspIleSerValAspPheIleGluValAlaIleValIleValIleValIle 240
DB 661 CCGTGTGTGACAAATCAGTAGTTTCATTGAAAGAAACCAACCTCCATGATGTTGTT 720
QY 241 LeuValHisCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleValIle 260
DB 721 CTAGTCACTGTTTACTGAGATCTCCGCTCCGCAACATGCTTACCTGCTACATCATG 780
QY 261 LysArgMetAspMetSerLeuAspGluIleValIleValIleValIleValIleValIle 280
DB 781 AAGAGATGAGCATGCTTATGATGAGCTTACAGATTTGTGAAGAAAGAAACCTACT 840
QY 281 IleSerProAsnPheAsnIleValIleValIleValIleValIleValIleValIleVal 300
DB 841 ATATCTCCAAATCTCAATTTCTGAGCAATCTGAGACTATGAGAGAAATTAAGAAC 900
QY 301 GlnThrGlyAlaSerGlyProIleSerIleValIleValIleValIleValIleValIleVal 320
DB 901 CAGACTGGAGCATCAGGCGCAAGAGCAAACTCACTGCTGCACTGAGAGAGCCAAAT 960
QY 321 GluProValProAlaValSerGluGlyIleValIleValIleValIleValIleValIleVal 340
DB 961 GAACCTGCTCCCTGCTCTCAGAGGAGTGGACAGAAAGCGAGCGCCCTCAGTCCACCC 1020
QY 341 CyAlaAspSerAlaHisSerGluAlaIleGlyIleValIleValIleValIleValIle 360
DB 1021 TGTGCCACTGCTGCTCACTCAGAGGAGCAGAGCAAAAGCCCTGATCCCGCACGCTG 1080
QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
DB 1081 CCCAGGCTGCCAGCTGACGCGCTGCTGCTGTTAGAGAGACAGCCGCTGATACAGCGCTC 1140
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleValIleValIleVal 400
DB 1141 AGTGGCTGACCTGCTCCGAGCAGCTGAGAGCAGCAATGAAGCTCAAGGCTTCTTC 1200
QY 401 SerLeuAspIleIleValSerValSerIleSerAlaSerMetAlaIleSerIleHisGlyPhe 420
DB 1201 TCTCTGAAATATCAAACTATTCATATTCAGGCGAGATGGCAGCATCTTACATGCTTC 1260
QY 421 SerSerSerGluAspAlaLeuGluIleValIleValIleValIleValIleValIleValIle 440

```

Db      1261 TCCTCATCGAAGATGCTTTGGATATCTACAAACCTTCCACTACTGTGATGGACCAAC 1320
Qy      441 LysLeuCyGlnPheSerProValGlnGluLeuSerGlnGlnThrProGluThrSerPro 460
Db      1321 AAGCTATGCGAGTCTCCCTCTGTTCAGAGAACTATCGAGACGACTCCCAAGACAGTCTCT 1380
Qy      461 AspLysGlnGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
Db      1381 GATAAAGAGAAAGCAGCATCTCCCAAGAACTCAGACCGCCAGGCTTCAGACAGCCAG 1440
Qy      481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db      1441 AGCAGAGATTGCAATGGTTCAGAACCCAGACAGAGTGGACCGCCCAAGAGTCTCTTTTA 1500
Qy      501 SerProLeuHisArgSerGlySerValGluAspArgTyrHisThrSerPheLeuPheGly 520
Db      1501 TCTTCACCTGCATGGAAGTGGAGCGTGAGAGCAATACACACAGACTTCCTTTTGGCC 1560
Qy      521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrHis 540
Db      1561 CTTTCCACAGCCAGCAGCAGCCTCAGAAAGTCTGCTGGCTGGCTTAAAGGCTGGCAC 1620
Qy      541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
Db      1621 TCGATATCTTGCCCCCAGACCTTACCCCTTCCCTGACACAGCAGCTGTATTTTGGC 1680
Qy      561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
Db      1681 ACGAGTCTCTCACCTTACTTCTGCTCAGCCATCTACGAGGCAAGTGCACATTACTCT 1740
Qy      581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
Db      1741 GCTTACAGCTGACAGCAGCTCCCACTTGGGAGACCAAGTCAATCTGTGGCGCAGGCGG 1800
Qy      601 GlnLysProSerAspAlaGlnAlaAspSerArgArgSerTyrHisGluGluSerProPheGlu 620
Db      1801 CAGAAAGCCAGATGACAGAGCTGACTCGCGCGAGCTGGAGTGAAGAGAGCCCTTTTAA 1860
Qy      621 LysGlnPheLysArgAspSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
Db      1861 AAGCATTTTAAACCGAAGAACTCCCAATGGAATTTTGAGAGAGCATCATGACAGAAAC 1920
Qy      641 ArgSerArgGlnGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db      1921 AGGTCAAGGAGAGAGCTGGGGAAGTGGGCACTGCTGAGCTTTTGGGCGACATGGA 1980
Qy      661 IleIleGluValSer 665
Db      1981 ATCATTCAGAGTCTCC 1995

```

RESULT 3

```

US-10-168-506-2
; Sequence 2, Application US/10168506
; Publication No. US20040053229A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: SUNDARSANAM, SUCHA
; APPLICANT: HILL, RON
; APPLICANT: FLANAGAN, PETER
; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
; FILE REFERENCE: 038602/1351
; CURRENT APPLICATION NUMBER: US/10/168, 506
; PRIOR APPLICATION NUMBER: 2002-06-21
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2732
; TYPE: DNA

```

ORGANISM: Homo sapiens
US-10-168-506-2

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0	3406.00	99.85%	99.70%	99.65%	13	2732	663	1	1	0	0

US-10-029-345a-109 (1-665) x US-10-168-506-2 (1-2732)

```

Qy      1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db      538 ATGGCCCATGAGATGATTGGAACTCAATATGTTACTGAGAGGTGGTGGCTGCTGGAA 597
Qy      21 SerGlyThrGlyLysValLeuLeuIleAspSerArgProPheValGluTyrArgThrSer 40
Db      598 AGTGAACGGAAAAAGTGGCTGCTAATGATTAAGCCGGCCATTTGTGGAATACATACATCC 657
Qy      41 HisIleLeuGlnAlaIleAsnIleAsnGlySerLysMetLysArgArgLeuGln 60
Db      658 CACATTTTGGAGCCATTAATATCAACTGCTCCAAAGCTTAAGAAACGAAAGGTTCACAG 717
Qy      61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db      718 GACAAAGTTAATTAACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 777
Qy      81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db      778 TGCAGTCAGAAAGTTGATTGATTACATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 837
Qy      101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db      838 GACTGTTTCTCAGTACTTCTGAGTAACTGGAGAAAGCTTCAACTCTGTTCACTG 897
Qy      121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyLeuGlyLysSer 140
Db      898 CTTCAGAGTGGGTTTGGCTGAGTTCTCTGCTGTTTCCCTGGCTCTGTGAAGAAATCC 957
Qy      141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db      958 ACTCTAGTCCCTTACTCTCATTTCTCAGCCTTGTCTTACTGTTGCCAACATTGGCCAACC 1017
Qy      161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuIle 180
Db      1018 CGAATTTCTCCCAATCTTTATCTTGGCTGCCAGCAGAGATGCTCTCAACAGAGAGCTGATG 1077
Qy      181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
Db      1078 CAGCAGAAATGGAAATGTATGTATGTTAAATGCCACAAATACCTGTCCAAAGCTGACCTTT 1137
Qy      201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
Db      1138 ATCCCCAGATCTCATTTCTGCTGGCTGCTGTGATGACACACTTTTGTGAAAAATTTTG 1197
Qy      221 ProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
Db      1198 CGGTGTTTGAACAAATTCAGTGAATTTCAATGAGAAACAAAGCTTCATGATGATGTGTT 1257
Qy      241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db      1258 CTAGTGACACTGTTTAGTGGGATCTCCCGCTCCGCCACCAATGCCATTCCTTACATCAG 1317
Qy      261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
Db      1318 AAGAGATGAGCAATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGACTTACT 1377
Qy      281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysLysIleLysAsn 300
Db      1378 ATATCTCCAAACTCATTTTCTGGGCGAACTCTGGAATAGAGAAAGATTAAAGAAC 1437

```



```

QY 201 ILeProGluSerHisPheLeuArgValProValIasnAspSerPheCysGluYsileLeu 220
DB 727 ATCCCGAGTCTCATTTCTCGGTGCTGTGATGACAGCTTTGTGGAAATTTTG 786
QY 221 ProTribLeuAspLysSerValAspPheIleGluYsAlaYsAlaSerHisGlyCysVal 240
DB 787 CCGTGGTGGCAAAATAGTAGATTTCATTGAGAAAGCAAAAGCTTCCATGAGATGTGT 846
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
DB 847 CTAGTGCACTGTTTAGTGGGATCTCCGCTCCGACCATGGCTATGAGATTAAGATC 906
QY 261 LysArgMetAspMetSerLeuAspGluAlaTYrArgPheValIleGluYsAlaYsAla 280
DB 907 AAGAGATGAGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 966
QY 281 ILeSerProAsnPheAsnPheLeuGlyGluLeuAspTYrGluYsAlaYsAlaYsAla 300
DB 967 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAGAAC 1026
QY 301 GluThrGlyAlaSerGlyProLysSerLysLeuYsLeuHisSleuGluYsProAsn 320
DB 1027 CAGACTGGAGACATCGAGGCGCAAAAGAGCAAACTCAAGGCTGCACTGGAGAAAGCAAA 1086
QY 321 GluProValProAlaValSerGluGlyGlyGlyLysSerGluThrProLeuSerPro 340
DB 1087 GAACCTGTCTCTGCTCTCAGAGGGTGGACAGAAAGCGAGCGCTTCAAGTCCACCC 1146
QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGluArgProValHisProAlaSerVal 360
DB 1147 TGTGCGGACTCTGCTACTCTCAGAGCGAGCAGCAAGCGCCGCTGATACAGCGCTTC 1206
QY 361 ProSerValProSerValGluProSerLeuLeuGluAspSerProLeuValGluAlaLeu 380
DB 1207 CCCAGCGGCCCAAGCGTGCAGCGCTGCTTGAAGAGACGCCGCTGATACAGCGCTTC 1266
QY 381 SerGlyLeuHisSleuSerAlaAspArgLeuGluAspSerAsnLysLeuYsArgSerPhe 400
DB 1267 AGTGGCTGCACCTGTCTCCGACAGAGCTGGAAGACAGCATATACCTCAAGCTTCTTC 1326
QY 401 SerLeuAspIleLysSerValSerTYrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
DB 1327 TCTCTGATATCAAAATCAGTTTCATTTACGCCACAGATGAGCATCTTTCATGCTTC 1386
QY 421 SerSerSerGluAspAlaLeuGluTYrTYrLysProSerThrThrLeuAspGlyThrAsn 440
DB 1387 TCTCTATCGAAGATGCTTTGGAATATCAAAACCTTCACTACTGATGGAGCAAC 1446
QY 441 LysLeuCysGluPheSerProValGluGluSerGluGluThrProGluThrSerPro 460
DB 1447 AAGCTATGCCAGTTCTCCCTGTTTCAAGCACTATCGAGAGCATCTCCGAAACCAATCTCT 1506
QY 461 AspLysGluGluAlaSerIleProLysLysLeuGluThrAlaArgProSerAspSerGln 480
DB 1507 GATTAAGAGAGAGAGCATCCCAAGAGCTGAGACCGCCAGGCTTCAAGACCAAG 1566
QY 481 SerLysValGluHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
DB 1567 AGCAAGCATTTGATTCGTGTAGAACCAAGACAGATGGACCGCCCAAGAGCTCTTTTAA 1626
QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTYrHisThrSerPheLeuPheGly 520
DB 1627 TCTCTCATGATCGAAGTGGAGCGTGGAGGACAAATTAACACCAAGGCTCTTTTGGAG 1686
QY 521 LeuSerHisSerGlnGlnHisSleuThrLysSerAlaGlyLeuGlyLysGlyTYrHis 540
DB 1687 CTTTCCACCGACAGCAGACCTCAAGAGCTGCTGGGCTTGAAGGCTGGGAC 1746
QY 541 SerAspIleLeuAlaProGluThrSerThrProSerLeuThrSerSerTYrTYrPheAla 560
DB 1747 TCGGATATCTTGGCCCCCAAGCTCTTACCTTCTGACCAAGCTGTGATTTTGGC 1806

```

```

QY 561 ThrGluSerSerHisPheTYrSerAlaSerAlaIleTYrGlyGlySerAlaSerTYrSer 580
DB 1807 ACAGAGTCTCTCACTTACTCTCTCCACCACTTACAGAGCACTGCGCAGTACTCT 1866
QY 581 AlaTYrSerCysSerGlnLeuProThrCysGlyAspGluAlaTYrSerValArgArg 600
DB 1867 GCTTACAGCTGAGCGCACTGCTCCCACTTGGGAGACCAAGCTTATCTGTGGCAGGCG 1926
QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerThrHisGluGluSerProPheGlu 620
DB 1927 CAGAGCCCAAGTGAAGAGCTGATCTCCGCGGAGCTGGGATGAGAGAGAGCCCTTGA 1986
QY 621 LysGlnPheLysArgSerCysGluMetGluPheGlyGluSerIleMetSerGluAsn 640
DB 1987 AAGCATTTTAAAGCAGAAAGCTGCAAAATGGAATTTGAGAGAGCATATGTCAGAGAAC 2046
QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
DB 2047 AGGTACAGGAGAGAGCTGGGGAAGTGGCAGTCACTTCTTTCGGCAGCATGAA 2106
QY 661 ILeIleGluValSer 665
DB 2107 ATCATTTAGGCTCTCC 2121

```

RESULT 5
US-09-964-277-1
Sequence 1, Application US/09964277
Patent No. US20020137170A1
GENERAL INFORMATION:
APPLICANT: Lucche, Ralf M.
TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.434
CURRENT APPLICATION NUMBER: US/09/964.277
CURRENT FILING DATE: 2001-09-25
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 3496
TYPE: DNA
ORGANISM: Homo sapiens
US-09-964-277-1

Alignment Scores:
Pred. No.: 0
Score: 3406.00
Length: 3496
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Matches: 663
Mismatch: 1
Query Match: 99.65%
Indels: 0
Gaps: 0

US-10-029-345a-109 (1-665) x US-09-964-277-1 (1-3496)

```

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 562 ATGGCCCAAGATATATGGAATCTCAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGGA 621
QY 21 SerGlyThrGluYsValLeuLeuIleAspSerArgProPheValGluTYrAsnThrSer 40
DB 622 AGTGAAGCGAAAAAGTCTCTTAATTGATAGCCGACCATTTGTGAAATACAAATCATCC 681
QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGln 60
DB 682 CACATTTTGAAGCCATTAATATCAATCTGCTCCAACTTATGAAAGCGAAGGTTGCAACG 741
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaYsHisLysValAspIleAsp 80
DB 742 GACAAAGTTTAATTACAGAGCTCATCCAGCATTCAGCAAAACATPAAGTTGACATTTGAT 801
QY 81 CysSerGlnLysValValValTYrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 802 TGCAGTCAGAAAGGTGTGATGTTACATCAAAAGCTCCCAAGATGTTGCTCTCTTTCA 861

```

Qy	101	AspCysPheLeuThrValIleuLeuGlyIysIleuGluIuIysSerPheAsnSerValHisIleu	120
Db	862	GACGTGTTTTCACACTGACTCTTCCTGGGTAACTGGAAAGAGCTTAACTCTGTTACCTGG	921
Qy	121	LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyIysSer	140
Db	922	CTTCACAGGTGGGTTGTCGAGGTCTCTCGTTGTTTCCCTGGGCTCTGTGAAGAAATCC	981
Qy	141	ThrIleuValProThrCysIleSerGlnProCysIleProValAlaAsnIleGlyProThr	160
Db	982	ACTCAGACCTCACTGACCTGCAATTCACGCTTGCTTACCTGTTGCAATGGGCAAC	1041
Qy	161	ArgIleIleuProAsnIleuTrpLeuGlyCysGlnAgaAspValIleuAsnIleGluIleu	180
Db	1042	CGAATCTCTCCCAATCTTATCTTGCGCCAGCGAAGATGTCCTAAACAAGAGGCTGATG	1101
Qy	181	GlnGlnAsnGlyIleGlyTrpValIleuAsnAlaSerTrpThrCysProIysProAspPhe	200
Db	1102	CAGCAGAAATGGGATTGGTTATGTGTAAATGCAACCAATCCTGTCCAAAGCTGACCTT	1161
Qy	201	IleProGluSerHisPheIleuArgValProValAsnAspSerPheCysGluIysIleu	220
Db	1162	ATCCCCGAGCTCAATTCCTCGTGTCGCTGGAATGACAGCTTTTGGAGAAATTTTGG	1221
Qy	221	ProTrpIleuAspIysSerValAspPheIleGlyIysAlaIysAlaSerAsnIlyCysVal	240
Db	1222	CCGGTGTTGGAACAATCAGTAGATTCTTGAAGAAAGAAAGCCTCCAAATGGAGTGT	1281
Qy	241	LeuValHisCysIleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTrpIleMet	260
Db	1282	CTAATGCACTGTTTACGTGGGATCTCCCGCTCCGCACCATGCTATGCGCTCAACTATG	1341
Qy	261	LysArgMetAspMetSerIleuAspGluAlaTrpArgPheValIysGluIysArgProThr	280
Db	1342	AAGAGGATGACATGCTCTTTAGATGAGACTTACAGATTGTGAAAGAAAGAAAGCACTACT	1401
Qy	281	IleSerProAsnPheAsnPheLeuGlyIleIleuLeuAspTrpGluIysIysIleIysAsn	300
Db	1402	ATATCTCCAACTTCAAATTTCTGGGCCAACTCCCTGACCTATGAAAGAAAGTTAAGAC	1461
Qy	301	GlnThrGlyAlaSerGlyProIysSerIysIleuIysIleuHisIleuGluIysProAsn	320
Db	1462	CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGGAGAGCCAAAT	1521
Qy	321	GluProValProAlaValSerGluGlyIleGlyIysSerGluThrProIysSerProPro	340
Db	1522	GAACTGTCCTCGTCTGTCAAGAGGTGACAGAAAGCAGAGCCCTCACTGATCACC	1581
Qy	341	CysAlaAspSerAlaThrSerGluAlaAlaGlyIleAspProValHisProAlaSerVal	360
Db	1582	TGTGCGCACTGTCACTCAGAGGCAACAGGACAAAGCCCGTGCATCCCGCCAGCGTG	1641
Qy	361	ProSerValProSerValGlnProSerIleuGluIysAspSerProIleuValGlnAlaIleu	380
Db	1642	CCCAAGCGGCCAAGGTGACAGCCGCTGCTTGAAGGACAGCCGCTGTATAGAGCGCTC	1701
Qy	381	SerGlyLeuHisIleuSerAlaAspArgLeuGluAspSerAsnIysIleuIysArgSerPhe	400
Db	1702	AGTGGGCTGCACCTGTCCGACAGACAGCTGAGAAAGCAACAATTAACCTCAAGGTTCTTC	1761
Qy	401	SerIleuAspIleuSerValSerTrpSerAlaSerMetAlaIleSerIleuHisGlyPhe	420
Db	1762	TCTCTGATATCAATCAGTTTCATATTCAGCCACACATGCGAGCATCTTCATCATGGCTTC	1821
Qy	421	SerSerSerGluAspAlaIleuGluTrpTrpIysProSerThrThrIleuAspGlyThrAsn	440
Db	1822	TCTTCATCAGAGAGGCTTTGGAAATCAACAACCTTCACATCTGTGATGGAGCAAC	1881
Qy	441	LysIleuCysGlnPheSerProValGlnIleuSerGluGlnThrProGluIysSerPro	460
Db	1882	AAGCATATCCAGTTCTCCCTGTTTACAGAACTATGAGAGACATCTCCGAAACCACTGCT	1941
Qy	461	AspIysGlnGluAlaSerIleProIysIleuGlnThrAlaArgProSerAspSerGln	480

Db	1942	GATAGAGGAGGAGGCACACATCCCAAGAGTGCAGACCCGACAGCCTTCAGACAGCCAG	2001
Qy	481	SerIysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu	500
Db	2002	AGCAAGGAGTGTGACTTGGGTGCAGAACAGACGACGAGTGGACCCGCCAGAGGCTCCCTTTTA	2061
Qy	501	SerProLeuHisAAsArgSerGlySerValGluIAspAsnTrpHisTrpSerPheLeuPheGly	520
Db	2062	TCCTCACTGCATCCAAAGTGGAGGAGGTGAGGACAAATTACACACACAGTTCCTTTTCGGC	2122
Qy	521	LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTrpHis	540
Db	2122	CTTTCCACACGACCCGACGACGACCTTCACGAGCTGTCTGGCCTGGGCTTTAAAGGCTGGAC	2183
Qy	541	SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTrpPheAla	560
Db	2182	TCGGATATCTTTGGCCCCCAGACCTTCACCCCTTCCTCCAGACACAGCTGTGTAATTTGGCC	2241
Qy	561	ThrGluSerSerHisAspMetTrsSerAlaSerAlaIleTrpGlyGlySerAlaSerTrsSer	580
Db	2242	ACAAAGTCCCTCACCTTCTACTCTGCTGCACGCCATCTACGAGGACAGGCCAGTACTCT	2301
Qy	581	AlaTrsSerCysSerSerGlnLeuProThrCysGlyValAspGlnValTrsSerValArgArg	600
Db	2302	GCTTACAGCTGCACGACGCTGCCACTTGGCGAAGCCAAAGCTTATCTGTGCGCAGCGG	2361
Qy	601	GlnLysProSerAspArgAlaAspSerArgTrsSerTrpHisGluGluSerProPheGlu	620
Db	2362	CAGAAAGCAAGTGCACAGAGCTGACTCCGGCGGAGCTGGCATGAAGAGAGCCCTTTGAA	2421
Qy	621	LysGlnPheLysValArgTrsSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn	640
Db	2422	AAGCAGTTTAAACCCCAAGAGCTGCCAAATGTGAATTGGAGAGACATCATGTACAGAAAC	2481
Qy	641	ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu	660
Db	2482	AGGTCACGGGAGAGAGCTGGGGAAAGTGGGACGTACGTCTAGCTTTTGGGCGACATGAA	2541
Qy	661	IleIleGluValSer	665
Db	2542	ATCATTTGAGGCTCTCC	2556
RESULT 6			
US-09-816-494-1			
Sequence 1, Application US/09816494			
Patent No. US20020034807A1			
GENERAL INFORMATION:			
APPLICANT: Meyers, Rachel A.			
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY			
TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR			
FILE REFERENCE: 10448-030002			
CURRENT APPLICATION NUMBER: US/09/816,494			
CURRENT FILING DATE: 2001-03-23/09/816,494			
PRIOR APPLICATION NUMBER: US 60/191,858			
PRIOR FILING DATE: 2000-03-24			
NUMBER OF SEQ ID NOS: 10			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 1			
LENGTH: 3544			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (589)...(2583)			
US-09-816-494-1			
Alignment Scores:			
Pred. No.: 0			
Score: 3406.00			
Percent Similarity: 99.85%			
Best Local Similarity: 99.70%			
Query Match: 99.65%			
Length: 3544			
Matches: 663			
Conservative: 1			
Mismatches: 1			
Indels: 0			

DB: 9 Gaps: 0

US-10-029-345a-109 (1-665) x US-09-816-494-1 (1-3544)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 Db 589 ATGGCCCATGAGATGATGGAACTCAATTTGTTACTGAGAGGTTGGTGGCTCGCTGGAA 648
 QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTrpAsnThrSer 40
 Db 649 AGTGGAAAGGAAAGGCTCTTATTTGATTAAGCCGGCATTTGTGGAAATACATATCATCC 708
 QY 41 HisIleuGlnuAlaIleAsnIleAsnGlySerLeuMetLeuArgGlnLeuGln 60
 Db 709 CACTTTTGGAGGCAATTAATATCACTGCTCCAACTTATAGCGAAGGTTGCCACAG 768
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
 Db 769 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATMAGTTGACATTGAT 828
 QY 81 CysSerGlnLysValValValIleTrpAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 Db 829 TGCAGTCAGAGGTTGTAGTTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 888
 QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
 Db 889 GACTGTTTCTCAGTACTTCTGGGTAACTGGAGAAAGCTTCAACTGTTCACCTG 948
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyLysSer 140
 Db 949 CTTCAGAGTGGGTTGCTGAGTCTCTGTTGTTCTGCTGGCTCTGTGAAGGAAATCC 1008
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 Db 1009 ACTTAGAGCTTCACTGATTTCTCAGCTTCTTACCTTGTGCGCAACATTTGGCGCAAC 1068
 QY 161 ArgIleLeuProAsnLeuTrpLeuGlyCysGlnArgAspValLeuAsnLysGluLeuIle 180
 Db 1069 CGAATTTCTCCCATCTTTATCTTGGCTGCGCAGCGAGATGTCCTCAACAGAGCTGATG 1128
 QY 181 GlnGlnAsnGlyIleGlyTrpValLeuAsnAlaSerTrpThrCysProLysProAspPhe 200
 Db 1129 CAGCAGATGGAGATGGTTATGTTAAATGCCAGCAATCTGTGCCAAAGCTGACCTTT 1188
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
 Db 1189 ATCCCGAGTCTCATTTCTCGGCTGCTGGAATGACAGCTTTTGGGAAATTTTG 1248
 QY 221 ProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
 Db 1249 CCGTGGTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAGGCTTCAATGTGTT 1308
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIle 260
 Db 1309 CTAGTGCACGTTTGTGCTGGATCTCCGCTGCCCAATCCCTATGCTCATCATG 1368
 QY 261 LysArgMetAspMetSerLeuAspGluAlaIleTrpPheValLysGluLysArgProThr 280
 Db 1369 AAGAGGATGACATGCTTTAGATGAAGCTTACAGATTCTGTGAAGAAAGAAAGACCTACT 1428
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTrpGluLysLysIleLysAsn 300
 Db 1429 ATATCTCCAACTTCATTTTCTGGCGCACTCTGAGCTATGAGAAAGATTAAAGAAC 1488
 QY 301 GlnThrGlyLysSerGlyProLysSerLysLeuLysLeuHisIleGlnLysPheProAsn 320
 Db 1489 CAGACTGAGCATCAGGCGCAAGCAAACTCAAGCTGCTGACCTTGAAGCGCAAT 1548
 QY 321 GluProValProAlaValSerGluGlyGluLysSerGluThrProLeuSerProPro 340
 Db 1549 GAACCTGTCCCTCTCTCAGGCGTGGACGAAAGCGAGCGCCCTCAGTCCAGCC 1608
 QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360

Db 1609 TGTGCCGACTCTGTACTTACAGAGCAGAGGACCAAGCCCGGTGATCCCGCAGCGGT 1668
 QY 361 ProSerValProSerValGlnProSerLeuLeuGluLysSerProLeuValGlnAlaLeu 380
 Db 1669 CCGAGCTGCGCAGAGGTGACAGCCGTGCTTTAAGAGCAGCCCGCTGTGATCAGGCGCTC 1728
 QY 381 SerGlyLeuHisIleSerSerAlaAspArgLeuGluAspSerLysLysLysArgSerPhe 400
 Db 1729 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGGCTTCTTC 1788
 QY 401 SerLeuAspIleLysSerSerValSerTrpSerAlaSerMetAlaAlaSerLeuHisIleGlyPhe 420
 Db 1789 TCTTGATATCAATCATGTTTCAATATTCAGCCAGCATGGCAGCATCTTATAGGCTTC 1848
 QY 421 SerSerSerGluAspAlaLeuGluTrpTrpLysProSerThrThrLeuAspGlyThrAsn 440
 Db 1849 TCTTCATCAGAGATGCTTTGGAATCTAACAACCTTCCACTACTGTGATGGAGCAAC 1908
 QY 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
 Db 1909 AAGCTATCCAGATTTCTCCCTGTTCAGGAATATGAGAGCAGACTCCCGAAACAGTCTT 1968
 QY 461 AspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 Db 1969 GATTAAGAGAGAGCCAGACTCCCAAGAACTGCAAGCCGACAGGCTTCAGACAGCAG 2028
 QY 481 SerLysArgLeuHisSerValArgHisSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 Db 2029 AGCAACGATTCATTCGTCGAGAACCGAGAGTGGACCGCCGAGAGTCCCTTTTA 2088
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTrpHisThrSerPheLeuPheGly 520
 Db 2089 TCTTCACGTCAGTCAAGTGGAGCGCTGAGGACATTAACAACAGCTTCTTTTCGCG 2148
 QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuLysGlyTrpHis 540
 Db 2149 CTTTCCACACACCGACAGCACCTCAAGAGTCTGTGGCTGGGCTTAAAGGCTGGAC 2208
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTrpPheAla 560
 Db 2209 TGGATATCTTGGCCCCCAGACCTTCACTTCCCTGACACAGAGCTGGATTTTGC 2268
 QY 561 ThrGlnSerSerHisPheTrpSerAlaSerAlaIleTrpGlyLysSerAlaSerLysSer 580
 Db 2269 ACAAGTCTCACAATTTCTTACTTGTCTGACCATCTACGAGGAGGCGCATTAATCT 2328
 QY 581 AlaTrpSerCysSerGlnLeuProThrCysGlyAspGlnValTrpSerValArgArg 600
 Db 2329 GCTTACAGCTGCAGCCAGCTGCCCACTTGGGAGACCAAGCTATTTCTGTGCGCAGGCG 2388
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGlnSerProPheGlu 620
 Db 2389 CAGAACCAAGTGAAGAGCTGACTGCGGCGGAGCTGGAGATGAAAGAGAGAGCCCTTTGAA 2448
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlnSerIleMetSerGluAsn 640
 Db 2449 AAGAGTTTAAACGCAAGCTGCCAAATGGAATTTGAGAGAGCATCATGTACAGAAAC 2508
 QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 Db 2509 AGGTACGCGAAGAGCTGGGAAAGTGGAGTCAAGTCACTTTTCTGGGAGCATGAA 2568
 QY 661 IleIleGluValSer 665
 Db 2569 ATCATTTGAGGTCTCC 2583

RESULT 7
 US-10-377-072-25

/ Sequence 25, Application US/10377072

/ Publication No. US20040009501A1

/ GENERAL INFORMATION:
 / APPLICANT: Millennium Pharmaceuticals Inc.

APPLICANT: Curtis, Rory A.J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Gluckemann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williams, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Miyoung
APPLICANT: Tsai, Fong-Ying
TITLE OF INVENTION: NOVEL 2869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38592, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: MPIO3-0180NAMIM
CURRENT APPLICATION NUMBER: US/10/377,072
CURRENT FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/895,860
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215,370
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/723,806
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/187,455
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/843,297
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,801
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: US 09/861,801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205,508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/816,494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/815,419
PRIOR FILING DATE: 2001-03-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Faeseq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 3544
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (589) ... (2586)
US-10-377-072-25
Alignment Scores:
Pred. No.: 0
Score: 3406.00
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 99.65%
Length: 3544
Matches: 663
Conservative: 1
Mismatches: 1
Indels: 0
Gaps: 0
US-10-029-345a-109 (1-665) x US-10-377-072-25 (1-3544)
QY 1 MetAlahieGluNetIleglyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 589 ATGGCCCATGAGATGATTGGAACCTCAATTTGTTACTGAGAGAGTTGGGCTGCTGCGAA 648
QY 21 SerGjLThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 649 AGTGAACGGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
QY 41 HisIleLeuGluValAlaIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 709 CACATTTTGAAGCATTAATATCACTGCTCCAGCTTATGAAGGAAAGTTGCAACG 768
QY 61 AspLyValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 769 GACAAAGTGTAAATTAACAGAGCTCATCCACATTCAGCGAAACATTAAGGTTGACATGAT 828
QY 81 CysSerGlnValValValValValValValValValValValValValValValValVal 100

DB 829 TGAAGTCAGAAAGTGTAGTATTACATCAAAAGCTCCAAAGATGTTCCCTCTCTCA 888
QY 101 AspCyPheLeuThrValLeuLeuGlyValLeuGluValSerPheAsnSerValHisLeu 120
DB 889 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTCCACTG 948
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyValSer 140
DB 949 CTTCAGAGTGGGTTTCTGATGTTCTCTGTTTCTCTGCTGCTCTGTTAGAGAAATCC 1008
QY 141 ThrLeuValProThrCysHisSerGlnProCysLeuProValAlaAsnIleglyProThr 160
DB 1009 ACTGATGCTCCATACCGATGATTTCTAGCCTTGTTACTGTTACCTGTTGCAACATGGCCAAAC 1068
QY 161 ArgIleLeuProAsnLeuValLeuGlyCysGlnArgAspValLeuAsnLeuLeuLeu 180
DB 1069 CGAATTCCTCCCAATCTTATCTTGGCTCCAGCGAGATGCTCTCAACAGAGCTGATG 1128
QY 181 GlnGlnAsnGlyIleglyThrValLeuAsnAlaSerThrCysProValProAspPhe 200
DB 1129 CAGCAGAAATGGAGTTGGTTATGTTAAATGCCAGCAATACCTGTCCAAAGCTTGACTTT 1188
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluValLeu 220
DB 1189 ATCCCGAGTCTCATTTCTGGGTGCTGATGAATGACAGCTTTGTGAGAAATTTTG 1248
QY 221 ProThrLeuAspLysSerValAspPheIleglyValAlaLeuAsnGlyCysVal 240
DB 1249 CCGTGGTTGGAACAATCACTAGATTTCAATTGAGAAAGCCCTCCATGATGTTGTT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaThrIleMet 260
DB 1309 CTAGGCACTGTATGAGTGGAGTCTCCGCTCCGACCATGCTATGCCCTACATCAATG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaValArgPheValIleglyValAspProThr 280
DB 1369 AAGAGATGACATGCTTTAATGAGAGCTTACAGATTTGTGAAGAAAGAAAGAAAGCTACT 1428
QY 281 HisSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspThrGluValHisLeuVal 300
DB 1429 ATATCTCCAACTTCAATTTCTGGGCAACTCTGAGTATGAGAAAGATTAAGAAC 1488
QY 301 GlnThrGlyAlaSerGlyProLysSerLysValLeuLeuHisLeuGluValProAsn 320
DB 1489 CAGACTGAGACATCAGGCGCAAGAGCAAACTCAAGCTCTGACCTGAGAGAGCCAAAT 1548
QY 321 GluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProPro 340
DB 1549 GAACCTGCTCCCTGCTCTCTCAAGAGGTGACAGAAAGGAGACAGCCCTCACTCAACCC 1608
QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
DB 1609 TGGCGCACTGCTGCTACCTCAGAGGCAAGCAAAAGCCGTTGATCCCGCAGAGT 1668
QY 361 ProSerValProSerValGlnProSerLeuLeuGluValAspSerProLeuValGlnAlaLeu 380
DB 1669 CCAAGGCTCCCAAGGCTGAGCGCTCGTTGAGAGAGACGCCGCTGTATACAGGCGCTC 1728
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuValSerPhe 400
DB 1729 AGTGGCTCACTGCTCCGCAAGAGCTGAGAGAGACAGAAATAGCTCAAGCTTCTTC 1788
QY 401 SerLeuAspLysSerValSerThrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
DB 1789 TCTCTGATATCAAAATCAATTTCAATTCAGCAGCATGCGACATCTTATACATGCTTC 1848
QY 421 SerSerSerGluAspAlaLeuGluValValValValValValValValValValVal 440
DB 1849 TCTCTATCAGAAAGTCTTGAATATCAAAACCTTCACTACTCTGATGAGCAAC 1908
QY 441 LysLeuCyGlnPheSerProValGlnLeuSerGluGlnThrProGluThrSerPro 460


```

Db      1909 AACCTATGCCAGTCTTCCCTGTTGACGAACTATCGACGAGCTCCCGAAACAGTCTT 1968
QY      461 AAsPlySGluGluA1Ser11eProLylySleuGlnThra1aArProSerAAsPserGln 480
Db      1969 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2028
QY      481 SerLySAgLeuH1aSerValArGThrSerSerSerGlyThra1aGlnArGserLeu 500
Db      2029 AGCAAGCCATTGCAATTCGGTCAGAACGACGACGACGACGACGACGACGACGACGACGAC 2088
QY      501 SerProLeuH1aAgSerSerGlySerValGluAspSerTrpH1aThrSerPheLeuPheGly 520
Db      2089 TCTCCACTGCATCCAAAGTGGAGGCTGTGAGAACAAATTACCAACAGCTTCTTTTCCGCG 2148
QY      521 LeuSerThrSerGlnGlnH1aSleuThrLySerValAGlyLeuGlyLeuLySGlyTrpH1a 540
Db      2149 CTTTCCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 2208
QY      541 SerAsp11eLeuA1aProGlnThrSerThrProSerLeuThrSerSerTrpLyPheA1a 560
Db      2209 TCGGATATCTTGGCCCCCCCCAGACTCTTACCCCTTCCCTGACACGACGACGACGACGACGAC 2268
QY      561 ThrGluSerSerH1aPheThySerVala1e1yrglyySera1aSerTySer 580
Db      2269 ACGAGATCTTCAACACTTCTACTCTGCTCTGACGACATTCAGAGGACAGGACGACGACGAC 2328
QY      581 AlaTySerCySerGlnLeuProThrCySerGlyAspGlnValTySerValArGArGArG 2388
Db      2329 GCTACAGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2448
QY      601 GlnLySProSerAAsPserA1aAsPserA1aSerA1a1e1yrglyySera1aSerTySer 620
Db      2389 CAGAAAGCCAAAGTGAAGAGCTGAGTCCGCGGCGAGCGATGAAGAGAGAGAGAGAGAGAGAG 2448
QY      621 LysGlnPheLySAgArGArGArGArGArGArGArGArGArGArGArGArGArGArGArGAr 2508
Db      2449 AAGCAGGTTTAAACCAAGAAAGCTGCAATGAAATTTGAGAGAGGATCATGCAAGAAC 2568
QY      641 ArgSerArgGluGluLeuGlyLySValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db      2509 AGGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2568
QY      661 Ile11eGluValSer 665
Db      2569 ATCATTTGAGGTCTCC 2583

```

US-10-425-114-26234
 ; Sequence 26234, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 26234
 ; LENGTH: 3625
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURES:
 ; OTHER INFORMATION: Clone ID: LIB4119-028-H6_FLI
 US-10-425-114-26234

Alignment Scores: 0 Length: 3625
 Pred. No.:

```

Score: 3406.00 Matches: 663
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 99.65%
DB: 13 Gaps: 0
US-10-029-345a-109 (1-665) x US-10-425-114-26234 (1-3625)
QY      1 MetAlaHisGluMet11eGlyThrGlnLeuValThraGluLeuValAlaLeuLeuGln 20
Db      692 ATGGCCCATGAGATGATTTGGAACTCAATATGTTACTGAGAGGTTGGCTGCTGCGA 751
QY      21 SerGlyThrGlyValLeuLeu1aAsPserA1aProPheValG1uTyraSerH1a 40
Db      752 AGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 811
QY      41 His11eLeuGluA1a1eAen11eAenCySerLySleuMetLyArGArGLeuGln 60
Db      812 CACATTTGGAGGCAATTAATATCAACTGCTCCAAAGCTTATGAAAGGAGGTTGCAAG 871
QY      61 AsPlyValLeu11eThrGluLeu11eGlnH1aSerA1a1e1yrglyyValAsp 80
Db      872 GACAAAGTGTAAATTACAGAGCTCATCCAGATTCAGGAAACATTAAGTTGAATGAT 931
QY      81 CySerGlnLySValValValTyraSpGlnSerSerGlnAspValAlaSerLeuSer 100
Db      932 TGCAGTCAGAGGTTGTAGTTTACATCAAAAGCTCCAAAGATGTGGCTCTCTCTTCA 991
QY      101 AspCyPheLeuThraValLeuLeuGlyLySleuGlnLySProPheAsPserVal1H1a 120
Db      992 GACTGTTTCTCAAGTACTTCTGGGTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1051
QY      121 Leu1aGlyGlyPheA1aGluPheSerArgCyPheProGlyLeuCySerGlyLyS 140
Db      1052 CTTGAGAGTGGGTTTGTGAGATTCCTGCTGTTTCCCTGAGGCTCTGTAAGAGAAATCC 1111
QY      141 ThrLeuValProThrCyS11eSerGlnProCySleuProValAlaAsn11eGlyProThr 160
Db      1112 ACTTACGCTCTTACCTGATTTCTCAAGCTTGTACTGTTGCTTCAAGCTTGTGCAAGCTTGG 1171
QY      161 Arg11eLeuProAsnLeuTyLeuGlyCyS1aArGArGArGArGArGArGArGArGArGAr 180
Db      1172 CGAATTTCTCCCAATCTTATCTTGGCTGCGACGAGAGATGCTTCAACAGAGAGTGAATG 1231
QY      181 GlnGlnAsnGly11eGlyTyraValLeuAsn1aSerTyThraCyPProLyProAsPhe 200
Db      1232 CAGCAAGATGGATTTGATATGTTAAATGCAACAAATACCTGTCCAAAGCTGACCTTT 1291
QY      201 11eProGluSerH1aPheLeuArGValProValaAsPserPheCyS1uLyS11eLeu 220
Db      1292 ATCCCGAGTCTCATTTCTGCGTGGCTGTGANTGACAGCTTTTGAGAGAAATTTTG 1351
QY      221 ProTrpLeuAsPlySerValaAsPhe11eGluLySAla1a1eAsPserGlyCyVal 240
Db      1352 CGGTGTTGAGCAAAATCGATGATTTCAATTGAAGAAAGCAAGCTTCATGATGTGTT 1411
QY      241 LeuValHisCySLeuA1aGly11eSerArGSerA1aThra11eA1a1eA1yry11eMet 260
Db      1412 CTAGTGCACTGTTTGTAGCTGGATCTCCCGGCTCCGACCACTGCTATCGCTTACATCAG 1471
QY      261 LysArGmetAsPmetSerLeuAsP1uA1aTyraPheVal1yS1uLySArGProThr 280
Db      1472 AAGAGATGAGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAGAGAGAGAGAGAGAG 1531
QY      281 11eSerProAsnPheAsnPheLeuGlyGlnLeuLeuAsPtyrGlnLyS1yry11eLyAsn 300
Db      1532 ATATCTCCAAATTTCTTGGGCAACTCTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1591
QY      301 GlnThraGlyA1aSerGlyProLySserLySleuLySleuGlnLySProAsn 320
Db      1592 CAGACTGAGACATCAAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGAGAGAGCAAAAT 1651
QY      321 GluProValProAlaValSerGluGlyGlnLySserGluThrProLeuSerProPro 340

```



```

Db      1652 GAACCTGCTCCCTGCTGCTCAGAGGCGTGAACAGAAAGCCAGACGCCCTCACTGACCC 1711
Qy      341  CysAlaAapSerAlaThrSerGluAlaAlaGlyAlaArgProValHisProAlaSerVal 360
Db      1712 TGTGCCACCTCTGCTACTCAGAGGCGAGGAGCAAAAGGCGCGGCTGATCCCGCAGCGTG 1771
Qy      361  ProSerValProSerValGlnProSerLeuGlnAapSerProLeuValGlnAlaLeu 380
Db      1772 CCCAGCTGCCCGCAGCGTCCGCTGTTAGAGACAGCCCGCTGGTACAGCGCTC 1831
Qy      381  SerGlyLeuHisLeuSerAlaAapArgLeuGlnAapSerValLeuValArgSerPhe 400
Db      1832 AGTGGCTGACCTGCTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCTTCTTC 1891
Qy      401  SerLeuAapLeuLeuSerValSerValSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
Db      1892 TCTCTGATATCAATTCAGTTTCAATTCAGCCAGCTGCGAGCATCTTACATGCGCTTC 1951
Qy      421  SerSerSerGlnAapAlaLeuGlnValTyrTyrLeuProSerThrThrLeuAapGlyThrAsn 440
Db      1952 TCTCTCATCAAGATGCTTGTGAATACTACAAACCTTCCACTCTGAGTGGAGCCAC 2011
Qy      441  LysLeuGlyGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
Db      2012 AACCTATGCCAGTTCTCCCTGTTCAAGGAATATCCGAGACAGACTCCGAAACCAAGCTCT 2071
Qy      461  AspLysGlnGlnAlaSerLysLeuValLeuGlnThrAlaArgProSerAapSerGln 480
Db      2072 GATTAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2131
Qy      481  SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db      2132 AGCAAGGATGATGCTTGTGTCAGAACAGAGCGTGGACCGCCGAGAGGCTCCCTTTTA 2191
Qy      501  SerProLeuHisArgSerGlySerValGlnAapSerValHisThrSerPheLeuPheGly 520
Db      2192 TCTCCACTGATGCAATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2251
Qy      521  LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyThrHis 540
Db      2252 CTTTCCACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2311
Qy      541  SerAspLeuLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrPyrPheAla 560
Db      2312 TCGGATATCTGGCCCGCCAGACCTTACCCCTTCCCTGACGAGCAGCTGGTATTTGGCC 2371
Qy      561  ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
Db      2372 ACAGAGCTCTCAACATTTCTACTGCTGCTCAGCCATCTACGAGGAGGAGGAGGAGGAG 2431
Qy      581  AlaTyrSerCysSerGlnLeuProThrCysGlyAapGlnValTyrSerValArgArgArg 600
Db      2432 GCCTTACAGCTGACGACGCTGCTCCACTTGGGAGAGCAAGCTATTTCTGGCCAGCGG 2491
Qy      601  GlnLysProSerSerPheArgAlaAapSerArgArgSerThrHisGlnGlnSerProPheGln 620
Db      2492 CAAGAAGCAAGTACAGAGCTGACTCCGCGGAGGAGCTGAGCATGAAGAGACCCCTTTGAA 2551
Qy      621  LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerLysMetSerGlnAsn 640
Db      2552 AAGCAGTTTAAAGCAGAGAGCTGCAATGGAATTTGGAGAGAGCATCATGTCAAGAAC 2611
Qy      641  ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
Db      2612 AGGTTCACGGGAGAGGCTGGGAGAAAGTGGCAGTCACTGCTTTTGGGCAACATGAA 2671
Qy      661  IleIleGlnValSer 665
Db      2672 ATCATTTAGGCTCTC 2686

```

RESULT 9
US-10-343-357-17

```

; Sequence 17, Application US/10343357
; Publication No. US20040058341A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.TOM
; APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
; APPLICANT: YAO, Monique G.; BUREFORD, Neil
; APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
; APPLICANT: LEE, Ernestine A.; HARALIA, April J.A.
; APPLICANT: LU, Dying Aina M.; TRIBOULEY, Catherine M.
; APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
; APPLICANT: YUE, Henry; WARREN, Bridget A.
; APPLICANT: NGUYEN, Danielle B.; CHAWLA, Narinder K.
; APPLICANT: KEARNEY, Liam
; TITLE OF INVENTION: PROTEIN PHOSPHATASES
; FILE REFERENCE: FI-0173 PCT
; CURRENT APPLICATION NUMBER: US/10/343,357
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: PCT/US01/23716
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,679
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/223,272
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/224,309
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 60/226,728
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/229,254
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: US 60/231,366
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 3766
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CB1
US-10-343-357-17

Alignment Scores:
Pred. No.: 0 Length: 3766
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 0
Query Match: 99.65% Indels: 0
DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-343-357-17 (1-3766)
Qy      1  MetaLHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
Db      538  ATGGCCCATGAGATGATTTGGAATCTCAATTTCTACTGAGAGAGTGGCTGCTGCGGAA 597
Qy      21  SerGlyThrGlnLysValLeuLeuIleAapSerArgProPheValGlnTyrAsnThrSer 40
Db      598  AGTGAACCGAAAGAAAGTCTGCTAATTGATTAGCGCGCCATTGTGGAAATACATATCATCC 657
Qy      41  HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGlnGln 60
Db      658  CACATTTTGAAGCCATTATATCAACTCTCCAACTTATGAAGCAAGTTGCAACAG 717
Qy      61  AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAapIleAap 80
Db      718  GACAAAGTGTATATTCAGAGCTCAATCCAGCATTCAGCGAAACATAGAGTTGACATTGAT 777
Qy      81  CysSerGlnLysValValValTyrAapGlnSerSerGlnAapValAlaSerLeuSerSer 100
Db      778  TGCAGTCAGAAAGTGTAGTTTACGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 837

```

```

QY 101 AspCyPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
Db 838 GACTGTTTCTCACTGACTGTTGAGTAACTGAGAGAGACTTCAACTGTTCACTG 897
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
Db 898 CTTGACAGGTGGGTGGTGGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAGGAAATCC 957
QY 141 ThrLeuValProThrCysHisSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 958 ACTTAAGTCCCTACCTGATTTCTCAAGCTGTGCTTACCTGTTCCCAACTTGGGCAACC 1017
QY 161 ArgIleLeuProAsnLeuLysLeuGlyCysGlnArgAspValLeuAsnLysLeuIle 180
Db 1018 CGAATTCCTCCCAATCTTATCTTGGCTGCACGAGATGCTCTCAACAAGAGCTGATG 1077
QY 181 GlnGlnAsnGlyIleGlyLysValLeuAsnAlaSerLysThrCysProLysProAspPhe 200
Db 1078 CAGCAGAAATGGAAATGGTTATGTTTAAATGCGACGAAATACCTGTCCAAAGCTGACTTT 1137
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
Db 1138 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGATGACACTTTTGTGAAGAAATTTTG 1197
QY 221 ProTrpLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCysVal 240
Db 1198 CCGTGGTGGACAAATCAGTACATTTCTTGAAGAAAGCAAAAGCCCAATGATGATGTT 1257
QY 241 LeuValHisCysLeuAlaGlyLysSerArgSerAlaThrIleAlaIleAlaTrpIleMet 260
Db 1258 CTAGTGCACTGTTTGTGGTGGATCTCCGCTCCGCCACATCGCTATCGCTTACATCATG 1317
QY 261 LysAlaGlyMetAspMetSerLeuAspGlyAlaAlaTrpPheValLysGluLysArgProThr 280
Db 1318 AAGAGAGATGACATGCTTTTGAAGACTTACAGATTTGTGAAGAAAGCAAAAGCACTACT 1377
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTrpGlyLysLysIleLysAsn 300
Db 1378 ATATCTCCAAACTTCAATTTCTGCGCCAACTCTGACATGAGAGAAAGTTTAAAGAC 1437
QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlnLysProAsn 320
Db 1438 CAGACTGGAGACATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTTGAGAAAGCAAAAT 1497
QY 321 GluProValProAlaValSerGluGlyGlnLysSerGluLysProLeuSerProPro 340
Db 1498 GAACCTGCTCCTCTGCTCTCAAGAGGTGACAGAAAGCAAGACGCCCCCTCAAGTCCACC 1557
QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
Db 1558 TGTGCCGACTGTGCTAAGCTCAGAGCAGACAGCAAGGACCCGCTGACGCGAGCGTG 1617
QY 361 ProSerValProSerValGlnProSerLeuGlnLysPheProLeuValGlnAlaLeu 380
Db 1618 CCAGAGTGTCCAGAGTGTGACGCTGTGTTTAAAGACAGCCGCTGTGACAGGCGCTC 1677
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
Db 1678 AGTGGCTGCACTGCTCGCAGACAGGCTGGAAGACAGCAATAGCTCAAGGCTTCTTC 1737
QY 401 SerLeuAspIleLysSerValSerLysSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
Db 1738 TCTCGGATATCAATCAATGTTTCAATATTCAGCCAGATGAGCATCTTATATGCTTC 1797
QY 421 SerSerSerGluAspAlaLeuGluLysLysProSerThrThrLeuAspGlyThrAsn 440
Db 1798 TCTTCATCAAGAGATGCTTGGAAATCTCAAAACCTTCACTACATCTGATGGAGCAAC 1857
QY 441 LysLeuCysGlnPheSerProValGlnLysLeuSerGluGlnThrProGluLysSerPro 460
Db 1858 AAGCTATGACAGTCTCCCTGTTCAGGAACATATGAGAGACACTCCGCAAAACAGCTCT 1917
QY 461 AspLysGluGlnLysSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480

```

```

Db 1918 GATAGAGAGAGAACCCAGACTCCCAAGAGCTGACAGACCGGACAGGCTTGACAGCAG 1977
QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db 1978 AGCAAGCTCATTCATTCGCTGAGAAACGACAGCATGGGACCGCCAGAGTCCCTTTTA 2037
QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTrpHisThrSerPheLeuPheGly 520
Db 2038 TCTTCACATGCAATCGAAGTGGAGCGGTGGAGACATTTACACACAGACTTCTTTTGGC 2097
QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysGlyTrpHis 540
Db 2098 CTTTCACACACCGACGACGACCTTCAAGAGTCTGAGCTGAGCTTGAAGGCTGGCAC 2157
QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTrpPheAla 560
Db 2158 TCGATATCTTGGCCCCCGACACTTACCCCTTCCCTGACAGCAGCGGTATTTTGC 2217
QY 561 ThrGlnSerSerHisPheTrpSerAlaSerAlaIleTrpGlyLysAlaSerTrpSer 580
Db 2218 ACAAGTCTCTCACTTCACTTCTGCTCAGCCATCTAAGAGGAGTCCAGTTACTCT 2277
QY 581 AlaTrpSerCysSerGlnLeuProThrCysGlyAspGlnValTrpSerValArgArg 600
Db 2278 GCTTACAGCTGACGACGCTGCCACTTGGGAGACCAAGTCTTATCTGTGCGCAGGCG 2337
QY 601 GlnLysProSerAspArgAlaAspSerArgSerTrpHisGluGluSerProPheGlu 620
Db 2338 CAAAGCTCAAGTGAAGAGCTGACTCGCGGAGCTGGCATGAAAGAACCCCTTTGAA 2397
QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyLysSerIleMetSerGluAsn 640
Db 2398 AAGCAGTTTAAACGACAAAGCTGCAAAATGAAATTTGAAGAGCAATCATGACAGAAAC 2457
QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db 2458 AGCTCAGCGGAAGAGCTGGGGAAAGTGGGACATGACTTATGCTTTTGGGACGATGAA 2517
QY 661 IleIleGluValSer 665
Db 2518 ATCATTTAGGTCTCC 2532

```

RESULT 10
 US-10-648-593-115
 / Sequence 115, Application US/10648593
 / Publication No. US20040106132A1
 / GENERAL INFORMATION:
 / APPLICANT: Bristol-Myers Squibb Company
 / TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
 / INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
 / FILE REFERENCE: D0273 NP
 / CURRENT APPLICATION NUMBER: US/10/648, 593
 / PRIOR APPLICATION NUMBER: 2003-08-26
 / PRIOR FILING DATE: 2002-08-27
 / NUMBER OF SEQ ID NOS: 557
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 115
 / LENGTH: 4790
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 US-10-648-593-115

Alignment Scores:
 Pred. No.: 0
 Score: 3406.00
 Percent Similarity: 99.85%
 Best Local Similarity: 99.70%
 Query Match: 17
 Length: 4790
 Matches: 663
 Conservative: 1
 Mismatches: 1
 Indels: 0
 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-648-593-115 (1-4790)

QY 1 MetAlHISGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 DB ATGGCCCATGAGATGATGGAACTCAAAATGTTACTGAGAGGTGGCTGCTGCGGAA 243
 QY 21 SerGlyThrGluIleValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 DB AGTGGAAAGGAAAGAGTGGCTCAATTTGATAGCGGCCATTTTGGAAATTCATACATCC 303
 QY 41 HisIleLeuGlnIleAlaIleAsnIleAsnCysSerIleLeuMetLysArgArgLeuGln 60
 DB CACATTTTGGAAAGCATTAATATCACTGCTCAAGCTTATGAAAGGAAAGTTGCAACAG 363
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
 DB AGCAAAAGTTAAATTCAGAGCTCATCCAGCATTCAGCCAAACATAAAGTTGACATGAT 423
 QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB TGCAGTCAGAAAGTTAGTTACGATCAAAAGCTCCCAAGATGTCCCTCTCTCTCA 483
 QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
 DB GACTGTTTCTCACTGTACTTGGGTAACCTGGAAAGAGCTTCAACTCTGTCACCTG 543
 QY 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
 DB CTTCGAGGTGGGTTTCTGAGTTCTCTCGTTGTTCTCTGCTCTGTCAGAAAGTATCC 603
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB ACTCTAGTCCCTACCTGCACTTCTCAGCTCTGCTACCTGTGTCCAAACATGGGCAACC 663
 QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
 DB CGAATTTCTTCCCAATCTTATCTTGGCTGCACGAGATGTCTCAACAAGAGCTGATG 723
 QY 181 GlnGlnAsnGlyTyrLeuGlyValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
 DB CAGCAAAATGGGATGTGTATGTGTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 783
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
 DB ATCCCGAGCTCATTTCTGCGTGTGCTGTGAATGACGCTTTGTGAAATTTTG 843
 QY 221 ProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
 DB CCTGTGTTGGAAACAAATCAGTATGATTTCTTGAAGAAAGCAAAAGCTTCAATGATGAT 903
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 DB CTAGTCACTGTTTACCTGGGATCTCCCGCTCGCACCATCTGCTATGCTCATCATG 963
 QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
 DB AAGAGGATGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAGAAAGCACTACT 1023
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLysAsn 300
 DB ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAAAGAAAGATTAAAGAC 1083
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGluLysProAsn 320
 DB CAGACTGAGACATCAAGGCGCAAAAGCAAACTCAACTGCTGCACTGGAGAAAGCAAAAT 1143
 QY 321 GluProValProAlaValSerGluGlyGlnLysSerGlnLysProLeuSerProPro 340
 DB GAACCTGCTCTCTGCTCAGAGGGGAGCAAAAGGCAAGAGCGCCCTCACTCAACCC 1203
 QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 DB TGTGCCGACGTCTGCTCTCAGAGGCAAGCAAGCAAAAGCGCTGTGATCCCGCACGCTG 1263

QY 361 ProSerValProSerValGlnProSerLeuGluAspSerProLeuValGlnAlaLeu 380
 DB CCAAGCTGCCCAAGCTGACCGCTGCTGTTAGAGAACAGCCGCTGGACAGGCGCTC 1323
 QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysValSerPhe 400
 DB AGTGGCTCACCTGTCCGACAGAGCTGGAAAGACAGCAATAGCTCAAGCTTCTTC 1383
 QY 401 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 DB TCTCTGATATCAAAATCAGTTTCAATTCAGCCAGCATGGAGCATCTTACATGCTTC 1443
 QY 421 SerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
 DB TCTCATCAGAAAGATGCTTGAATATCTCAAACTTCACTACTGTGATGGAGCAAC 1503
 QY 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
 DB AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATCGGAGCAAGCTCCGAAAGCAATCCT 1563
 QY 461 AspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 DB GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGACAGCCGAGGCTTACAGACAGCAG 1623
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 DB ACGAAGCGATTCATTCGGTCAAGAACACAGACAGAGGAGCCGCGCAAGAGTCCCTTTTA 1683
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
 DB TCTCCACTCATCGAAGTGGAGCGTGGAGAGCAATTACACACAGCTTCTTTTCGGC 1743
 QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheHis 540
 DB CTTTCACAGCAGCAGCAGACCTCAGAGATCTGCTGGCTGGGCTTAAAGGCTGGCAC 1803
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPyrPheAla 560
 DB TGGATATCTGGAGCCCGCAGACCTTACCCCTCTCTACCAAGAGCTGTATTTTGGC 1863
 QY 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysAlaSerTyrSer 580
 DB ACGAGTCTCTCACCTTACTCTGCTCAGCATTTACAGAGCAGTCCAGTTACTCT 1923
 QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
 DB GCTACAGCTGCAGCAGCTGCCCACTTCCGAGAACCAAGTCTATTCTGCGCAGAGCGG 1983
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGluGluSerProPheGlu 620
 DB CAGAAAGCCAAAGTACAGAGCTACCTCGCGGAGCTGGCAATGAAAGAGGCCCTTTGAA 2043
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
 DB AAGCACTTTAAACGCAAGAGCTGCCAAATGGAATTTGGAGAGCATCATGTCAGAAC 2103
 QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 DB AGGTACCGGAAAGAGCTGGGAAAGTGGGACATGACTTCTTCTTGGGAGCATGAA 2163
 QY 661 IleIleGluValSer 665
 DB ATCATTTGAGGTCTCC 2178

RESULT 11
 US-10-094-749-673
 ; Sequence 673, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI

```

      0
      3399.00
      99.70%
      99.55%
      99.44%
      16

```

US-10-094-749-673 (1-2102)

QY	1	MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGlu	20
Db	56	ATGGCCCATGACATGATTGGAACTCAATTGTCTAGAGGTTGGGGCTCTGGTGGAA	111
QY	21	SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnArgAsnThrSer	40
Db	116	AGTGAACCGCAAAAAGTCTGCTAATTGATTAGCCGGCCATTGTGGAAATACATCATCC	176
QY	41	HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGlnGln	60
Db	176	CACATTTTGGAAAGCATTTATATATACATGCTGCACACTTATACAGCAAGGTTCGACACG	236
QY	61	AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp	80
Db	236	GACAAAGGTTAATTATACAGAGCTCATCCAGCATTCAGCAACATTAAGGTGGAATTGAT	295
QY	81	CysSerGlnLysValValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer	100
Db	296	TGCAGTCAGAAAGGTTTGAATTGACGATCAAACTCCAAAGTGTGCTCTCTCTCTCA	355
QY	101	AspCysPheLeuThrValLeuLeuGluGlyLysLeuGlnLysSerPheAsnSerValHisLeu	120
Db	356	GACGTTTTCTCATCTGTACTTCTGGGTGAACCTGGGAAGAACCTTCACCTCTGTCACTGG	415
QY	121	LeuAlaGlyValPheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlyLysSer	140
Db	416	CTTGCAGAGTGGGTTGCTGAGTTCTCTGTTGTTTCCCTGGCCTCTGTGCAAGAAATATCC	475
QY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	476	ACCTTAGTCCCTACCTGACATTTCTCAGGCTTGGCTTACTGTTCACAACTTTGGGCCAACCC	535

Qy	161	ArgIleLeuPProAsnLeuTYrLeuGlyCysGlnAArgPValLeuAsnLeuLeuIle	180
Db	536	CPAAITCTCTCCCAATCTTTATCTATCTGGCTCCAGCGAGATGTCCTCCAAACAGAGCTGATG	595
Qy	181	GlnGlnAsnGlyIleGlyTYrValLeuAsnIleSerTYrThrCysProIysProAspPhe	200
Db	596	CAGCAGAAATGGATGCTTATGCTTAAATGCAAGCAATACCTGTCCAAAGCCTGACTTT	655
Qy	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIysIleLeu	220
Db	656	ATCCCGAGTCTCATTTCTCGGTGTGCTGTGAATACAGCTTTTGAGAAATATTTG	715
Qy	221	ProTyrLeuAspLysSerValAspPheIleGlyIysValAlaValAsnLeuIleCysVal	240
Db	716	CCGTGGTTTGCAAAATCAGTAGATTATTGAGAAAGCAAAAGCCCTCCAAATGAAAGTCTT	775
Qy	241	LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleTYrIleMet	260
Db	776	CTATGGCACTGTTTACGTGGAGATCTCCCGCTCCGACCATGCTATGCGCTACATCATG	835
Qy	261	IysArgMetAspMetSerLeuAspGlnAlaTYrArgPheValLysGlnLysArgProThr	280
Db	836	AAGAGAGAGAGCATGCTCTTATAGTAAGCTTACAGATTGTGAAAGAAAAGAACCTCACT	895
Qy	281	IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTYrGlnLysValIleLysAsn	300
Db	896	ATATCTCCAAACTTCATTTTCTGGCCCAACTCTCGAGACTATGAGAAAGATTAAAGAAC	955
Qy	301	GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlnLysProAsn	320
Db	956	CAGACTGACATCAAGGCGCAAGAGCAAACTCAAGCTGTGCACCTGGAGAGCCCAAT	1015
Qy	321	GlnProValProAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerProPro	340
Db	1016	GAACCTGTCTCTGCTGTCTCAGAGGGTGAGACAAAAGCAGAGGCCCTCACTCCACCC	1075
Qy	341	CysAlaAspSerAlaThrSerGlnAlaIleGlyGlnArgProValHisProAlaSerVal	360
Db	1076	TGTGGCCACTGTGCTACCTCAAGAGCAGCAGCAAAAGCCCTGTGCATCCGCCAGCGTG	1135
Qy	361	ProSerValProSerValGlnProSerLeuGlnLysAspSerProLeuValGlnAlaLeu	380
Db	1136	CCAGGGTCCCAAGCGTGACCCCTGCTGTAGAGACAGCCCGCTGTACAGCGCTTC	1195
Qy	381	SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe	400
Db	1196	AGTGGGCTGCACCTGTCCGACAGAGGCTGGAAAGACAGCAATTAACCTCAACCTTCTTC	1255
Qy	401	SerLeuAspIleLysSerValSerTYrSerAlaSerMetAlaIleSerLeuHisGlyPhe	420
Db	1256	TCTCTGATATCAATATAGTTTATATTTCAGCACACATGCGAGCATCTTACAGGCTTC	1315
Qy	421	SerSerSerGlnAspAlaLeuGlnTYrTYrLysProSerThrThrLeuAspGlyTYrAsn	440
Db	1316	TCCCATCACAAAATGCTTTGGAAATCAACAACTTCACTACTCTGATGGAGCAACAC	1375
Qy	441	LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro	460
Db	1376	AAGCTATGCCAGTTCTCCCTCTTTACGAACTTATCGAGAGAGACTCCCGAAACAACTCT	1435
Qy	461	AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerArgSerGln	480
Db	1436	GATTAAGAGAGAGAGCAGATCCCAAGAAAGCTGCAGACTCCAGGCCCTTACAGACAGCAG	1495
Qy	481	SerLysArgLeuHisSerValArgThrSerSerSerGlyTYrAlaGlnArgSerLeuLeu	500
Db	1496	AGCAACCATGTCATTCCTGTCAGAAACAGACAGCGAGGACCGCCCAAGAGGCTCTTTTA	1555
Qy	501	SerProLeuHisArgSerGlySerValGlnAspAsnTYrHisThrSerPheLeuPheLys	520
Db	1556	TCTTCACATGCATCAAGATGGAGGCTGTGGAGCAATTACACACAGACTCTCTTTTGACC	1615
Qy	521	LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyIleuLysGlyTYrHis	540

Db 1616 CTTTCCACCCAGCAGCAGCCTCAGCAAGCTGCTGCGCTTGAAGGCTGGCAC 1675
Qy 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpYrPheAla 560
Db 1676 TCGGATATCTTGAGCCCGCCAGACCTCTACCCCTTCCCTGACAGAGCTGGATTGTC 1735
Qy 561 ThrGlnSerSerIshPheYrSerAlaSerAlaIleYrGlyGlySerAlaSerYrSer 580
Db 1736 ACGAGCTCTCCACACTTCTACTCTGCTCAGCATCTACGAGGAGAGCTGCACTTCT 1795
Qy 581 AlaYrSerCySerGlnLeuProThrCyGlyAspGlnValYrSerValAlaGlyArg 600
Db 1796 GCTTACAGCTGCGACCCAGCTGCCCACTTGGGAGACCAAGCTATTCTGCGCCAGCG 1855
Qy 601 GlnLysProSerAspAlaAlaAspSerArgArgSerTrpHisGlnLysProPheGlu 620
Db 1856 CAGAGCCAAAGTACAGAGCTGACTCGCGCGAGCTGGCATGAGAGAGCCCTTGA 1915
Qy 621 LysGlnPheLysArgArgSerCyGlnMetGlnPheGlyGlnSerIleMetSerGluAsn 640
Db 1916 AACGAGTTTAAAGCAGAGAGCTGCCAAATGGAAATGGAGAGAGCATCATGTCAGAAC 1975
Qy 641 ArgSerArgGlnGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db 1976 AGCTACGGGAAGAGCTGGGGAAGTGGCAGTCACTGCTTTTGGGCAAGATCGAA 2035
Qy 661 IleIleGlnValSer 665
Db 2036 ATCATTGAGGTCTCC 2050
RESULT 12
US-10-296-115-520 ; Sequence 520, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hysed Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 520
; LENGTH: 2966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2966)
; OTHER INFORMATION: n = a, t, c or g
US-10-296-115-520
Alignment Scores:
Pred. No.: 0 Length: 2966
Score: 3398.00 Matches: 662
Percent Similarity: 99.70% Conservative: 1
Best Local Similarity: 99.55% Mismatches: 2
Query Match: 99.41% Indels: 0
DB: 13 Gaps: 0
US-10-029-345a-109 (1-665) x US-10-296-115-520 (1-2966)
Qy 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
Db 23 ATGGCCCATAGATGATTGGAACTCAATATTGTTACTGAGAGGGGGCTGCTGGAGAA 82
Qy 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnYrAsnThrSer 40
Db 83 AGTGGAAAGGAAAGTGTCTGTTAATTGATAGCCGGGCAATTGTGGAAATACATATCTCC 142

Qy 41 HisIleLeuGlnAlaIleAsnIleAsnCySerIleValMetLysArgLeuGln 60
Db 143 CACATTTGGAAAGCCATTATATCACTGCTCCACCTTATGAAAGGAGTTGCACAG 202
Qy 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 203 GACAAAGTGTATTATACAGAGCTCATCCAGCAATTCAGCGAAACATAGGTGACATTGAT 262
Qy 81 CySerGlnLysValValValYrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 263 TGCAGTCAGAAAGGTTGATTATGATCAAGCTCCCAAGATGTTCCCTCTCTCTTCA 322
Qy 101 AspCyPheLeuThrAlaLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 323 GACTGTTTCTCATCTGACTTCTGGGTAAACTGGAAGAAGCTTCAACTGTTCACCTG 382
Qy 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCyPheProGlyLeuCyGlnGlyLysSer 140
Db 383 CTTCGAGGTGGGTTCCTGACTGCTCTCTGTTTCCCTGCTCTGTAAGGAAATCC 442
Qy 141 ThrLeuValProThrCySileSerGlnProCySleuProValAlaAsnIleGlyProThr 160
Db 443 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACTGTTGCCAAACATTTGGGCAACC 502
Qy 161 ArgIleLeuProAsnLeuYrLeuGlyCyGlnArgAspValLeuAsnLysGlnLeuIle 180
Db 503 CGAATTTCTCCCAATTTTATCTTGGCTGCCAGCGAAGTCTCTCAACAGAGGCTGATG 562
Qy 181 GlnGlnAsnGlyIleGlyYrValLeuAsnAlaSerYrThrCySProLysProAspPhe 200
Db 563 CAGCAGAAATGGATTTGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 622
Qy 201 IleProGlnSerIshPheLeuArgValProValAsnAspSerPheCyGlnLysIleLeu 220
Db 623 ATCCCGAGTCTCATTTCTGCGTGGCTGCGTAAGAGACAGCTTTTGTGGAAAAATTTTG 682
Qy 221 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAsnGlyCySVal 240
Db 683 CCGTGTGTGACAAACAGTATGATTTTCATTGAGAAAGCAAAACCTCCAAATGATGTGT 742
Qy 241 LeuValHisCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaYrIleMet 260
Db 743 CTAGTCACTGTTTACTGCTGGATCTCCGCTCCGCAACATCGCTATCGCTTATCATG 802
Qy 261 LysArgMetAspMetSerLeuAspGlnAlaYrArgPheValLysGlnLysArgProThr 280
Db 803 AAGAGATGAGACATGCTTTAATGATGAAGCTTACAGATTGTGAAAAAAGAAAGCACTACT 862
Qy 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTrpGlnLysIleYrAsn 300
Db 863 ATATCTCCAAACTCAATTTTCTGGGCAACCTCTGACTATGAGAAAGATTAAGAAC 922
Qy 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlnLysProAsn 320
Db 923 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAACTGCTGCACTGAGAGAGCCAAAT 982
Qy 321 GluProValProAlaValSerGlyGlyGlnLysSerGlnThrProLeuSerProPro 340
Db 983 GAACTGCTCCCTGCTCTCAGAGGGTGAAGAAAGAGAGAGCGCCCTCACTCCACACC 1042
Qy 341 CySAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
Db 1043 TGTGCGACTGTGCTACCTCAGAGGAGAGAGCAAAAGCCCTGTGATCCCGCACCGTG 1102
Qy 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
Db 1103 CCGAGCTGCCAGCTGAGCCGTGCTGTATGAGAGACAGCCCGTGGTACAGGGGCTC 1162
Qy 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLysArgSerPhe 400
Db 1163 AGTGGGCTGCACTGCTCCGACAGCAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTCT 1222

```

QY 401 SerLeuAsp11eLysSerValSerTySerAlaSerMetAlaAlaSerLeuH1sglypHe 420
Db 1223 TCTCTGGATATCAATCAAGTTTCATATTCACGCGACATGGCAGCATCTTTACATGGCTTC 1282
QY 421 SerSerSerGluAspAlaLeuGluValTyTyrPheSerThrThrLeuAspGlyThrAsn 440
Db 1283 TCCCTATCAGAAAGATGCTTGGAAATACAAACCTTCCACTCTCTGAGTGGACCAAC 1342
QY 441 LysLeuCySGlnPheSerProValGlnGluLeuSerGlnGlnThrProGlnThrSerPro 460
Db 1343 AAGCTATGCCAGTCTTCCCTGTTTCAGAACTATCGAGAGCAGACTCCGAAACAGTCTCT 1402
QY 461 AspySGlnGluAlaSer11eProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
Db 1403 GATTAAGAGAGAGGACCATCCCAAGAACTGCAACCGCGAGCTTCCAGACAGCCAG 1462
QY 481 SerLysArgLeuH1sSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db 1463 AGCAAGCGATTGCTTGGTTCAGAACCAAGCAGAGGCGACCGCCAGAGGCTCCCTTTA 1522
QY 501 SerProLeuH1sArgSerGlySerValGluAspAsnTyrH1sThrSerPheLeuPheGly 520
Db 1523 TCTTCACTGATGAGAGTGGAGCGGAGGAGCAATTCACACAGCTTCTTTTCGCGC 1582
QY 521 LeuSerThrSerGlnGlnH1sLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPhe 540
Db 1583 CTTTCCACAGCAGCAGAGCAGCTTCAAGAACTCTGCTGGCTGGCGCTTAAGGCTGGCAGC 1642
QY 541 SerAsp11eLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrThrPheAla 560
Db 1643 TCGGATATCTTGGCCCCCAGACCTTACCTTCCCTGACACAGCAGCTGGATTTTGGC 1702
QY 561 ThrGlnSerSerH1sPheTySerValSerAlaLeuTyTyrGlyLeuSerAlaSerTySer 580
Db 1703 ACAGAGTCTCTACACTTCTGCTCTGCTCAGCAATCTACGAGAGCGAGTCCAGTACTCT 1762
QY 581 AlaTySerCySGlnGlnLeuProThrCyGlyAspGlnValTySerValArgArgArg 600
Db 1763 GCTTACAGCTGACAGCTGCTGCCACTTGGCGAGACCAAGTCAATCTCTGCGACGGCGG 1822
QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerThrH1sGlnGluSerProPheGln 620
Db 1823 CAGAAAGCAAGTACAGAGCTGACTCGCGCGAGTGGCAGTGAAGAAAGCCCTTTGAA 1882
QY 621 LysGlnPheLysArgArgSerCySGlnMetGluPheGlyGluSer11eMetSerGluAsn 640
Db 1883 AAGCAATTAAAGCGAAGGCTGCCAATGAAATTTGAGAGAGCATCATCTCAGAGAAC 1942
QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db 1943 AGGTCAACGGGAAGAGCTGGGGAAGTGGGCACTCAGTCTTTTGGGACAGCATGAA 2002
QY 661 Ile11eGluValSer 665
Db 2003 ATCATTTAGAGTCTCC 2017

```

RESULT 13

```

US-10-072-012-257
; Sequence 257, Application US/10072012
; Publication No. US20040033493A1

```

GENERAL INFORMATION:

```

; APPLICANT: Tcherny, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zetshueen, Bryan
; APPLICANT: Patcurajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie

```

```

; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimír Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Riegey, Denise M.
; APPLICANT: Riegey, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR APPLICATION NUMBER: 2002-01-31
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 257
; LENGTH: 2071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-012-257
Alignment Scores:
Pred. No.: 0
Score: 3379.50
Percent Similarity: 99.40%
Best Local Similarity: 99.25%
Query Match: 98.87%
Dbs: 13
Gaps: 1
US-10-029-345a-109 (1-665) x US-10-072-012-257 (1-2071)
QY 1 MetAlaH1sGluMet11eGlyThrGln11eValThrGluArgLeuValAlaLeuGlu 20
Db 61 ATGGCCCATGAGATGATTTGAACTCAATTTGTTACTAGAGCTTGGTGGCTGGTGGAA 120
QY 21 SerGlyThrGluLysValLeuLeu11eAspSerArgProPheValGluTyrAsnThrSer 40
Db 121 AGTGGACGGAAGAAAGTCTCTCTAATGATGAGCGGCAATTTGTGGAAATACATATCC 180
QY 41 His11eGluGluAla11eAsn11eAsnCySerLysLeuMetLysArgArgLeuGlnGln 60
Db 181 CACATTTTGAAGCATTAATATCACTGCTCAACCTTATGAGCGAAGGTTGCAACAG 240
QY 61 AspyValLeu11eThrGluLeu11eGlnH1sSerAlaLysH1sLysValAsp11eAsp 80
Db 241 GACAAAGTTAATATACAGAGCTCATTCAGCATTCAGCGAACAATAGTTGACATTGAT 300
QY 81 CySerGlnLysValAlaValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 301 TCGAGTCAGAAAGTGTAGTTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 360

```

QY 101 AapCyPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheLeuSerValHisLeu 120
Db 361 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAAAGAGCTTCACTGTTCACTG 420
QY 121 LeuAlaGlyLysPheAlaGlyPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
Db 421 CTTCAGAGTGGGTTTCTGAGTTCTCTGTTGTTTCCCTGCTCTGTGAGAAATTC 480
QY 141 ThrLeuValProThrCysHisSerGlnProCysLeuProValAlaAsnHisGlyProThr 160
Db 481 ACTGATGCTCCCTACCTGCACTTCTCAAGCTTGTCTTACCTGTGCAACATTTGGCCAC 540
QY 161 ArgHisLeuProAsnLeuThrLeuGlyCysGlnArgAspValLeuLeuGlyLysLeu 180
Db 541 CGAATCTTCCCAATCTTATCTTGGCTGCACGAAATGCTCTCAACAAAGAGCTGATG 600
QY 181 GlnGlyAsnGlyHisGlyThrValLeuAsnHisSerThrCysProLysProAspPhe 200
Db 601 CAGCAGAAATGGGATTTGGTTATGTGTTAAATGCCAGCAATCTGTCCAAAGCTGACTTT 660
QY 201 HisProGlySerHisPheLeuArgValProValAsnAspSerPheCysGlyLysLeu 220
Db 661 ATCCCGAGTCTCATTTCTCGCGGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 720
QY 221 ProThrLeuAspLysSerValAspPheHisGlyLysAlaLysSerAsnGlyCysVal 240
Db 721 CCTGTGTGACAAATCAGTAGATTTCTATGAGAAACAAAGCTCTCAATGATGTTG 780
QY 241 LeuValHisCysLeuAlaGlyLysSerArgSerAlaThrHisLeuAlaHisLeuMet 260
Db 781 CTATGTCACCTTTTAACTGAGGATCTCCGCTCCGCCAATCGCTATGCTTACATCATG 840
QY 261 LysArgMetAspMetSerLeuAspGlyAlaThrArgPheValHisGlyLysArgProThr 280
Db 841 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 900
QY 281 HisSerProAsnPheAsnPheLeuGlyLysLeuLeuAspGlyGlyLysHisLeuAsn 300
Db 901 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTAAGAAC 960
QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlyLysProAsn 320
Db 961 CAGACTGAGACATCAGGGCCAAAGAGCAATCAAGCTGTGACCTGTGAGAAAGCAAT 1020
QY 321 GluProValProAlaValSerGlyGlyLysSerGlyLysProLeuSerProPro 340
Db 1021 GAACCTGCTCCTGCTGCTCAGAGGGTGAAGAGAAAGCAAGCGCCCTCAATCCACC 1080
QY 341 CysAlaAspSerAlaThrSerGlyAlaHisGlyLysArgProValHisProAlaSerVal 360
Db 1081 TGTGCCACTCTCTCTACCTAGAGGCGAGCAAGAAAGCCCGTGCATCCCGCC----- 1124
QY 361 ProSerValProSerValGlnProSerLeuLeuGlyLysAspSerProLeuValGlnHisLeu 380
Db 1135 ---AGCTTACCCAGCGTGCAGCGCTGCTGTTAGAGACAGCCCGCTGTGACAGCGCTC 1191
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlyLysSerAsnLysLeuLysArgSerPhe 400
Db 1192 AGTGGGCTGACCTGCTCGCAGAGAGGCTGGAAGACAGCAATAGCTCAAGCGTTCTTC 1251
QY 401 SerLeuAspHisLeuSerValSerThrSerAlaSerMetAlaAsnLeuHisGlyPhe 420
Db 1252 TCTCTGATATCAAAATCAGTTTCAATATTCAGCCAGCATGTGCAGATCTTACATGGCTTC 1311
QY 421 SerSerSerGlnAspAlaLeuGlyLysThrLysProSerThrHisLeuAspGlyHisAsn 440
Db 1312 TCTCTCATCAAAAGATGCTTGGATCTTACAAACCTTCCATCTCTGATGGGACCAAC 1371
QY 441 LysLeuCysGlnPheSerProValGlnGlyLysSerGlyLysHisProGlyLysSerPro 460
Db 1372 AAGTATGACAGTCTCCCTGTTGAGAACTATCGAGAGCACTCCCGAAACAGTCTCT 1431

QY 461 AspLysGlyLysAlaSerHisLeuProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
Db 1432 GATTAAGAGAGAACCCAGCATTCCTCCAAAGAGCTGCAACCGCCAGGCTTCAAGACCCAG 1491
QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db 1492 AGCAAGCATTTGATTCGGTCAAGAACACAGAGAGTGGACCGCCAGAGGTCCCTTTTA 1551
QY 501 SerProLeuHisArgSerGlySerValGluAspAsnThrHisThrSerPheLeuPheGly 520
Db 1552 TCTCCATCTCATCGAAGTGGAGGCTGGAGAGAAATTAACCAACAGCTTCTTTTCGGC 1611
QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysLeuGlyLysPhe 540
Db 1612 CTTCACCAAGCCAGAGCACTTCAAGAACTCTGCTGGCTGGCTTAAAGGCTGGAC 1671
QY 541 SerAspHisLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrPheAla 560
Db 1672 TGGATATCTTGGCCCCCAGACCTTACCCCTTCCAGACAGAGCTGTATTTGGC 1731
QY 561 ThrGlySerSerHisPheThrSerAlaSerAlaHisLeuGlyLysSerAlaSerThr 580
Db 1732 ACAGAGTCTCTCACTTCTTACTTCTGCTCAGCATCTACAGAGGAGTGCAGTACTCT 1791
QY 581 AlaThrSerCysSerGlnLeuProThrCysGlyAspGlnValThrSerValArgArgArg 600
Db 1792 GCTTACAGCTGAGCCAGCTGCTGCTGCTGAGAGCAAGTCTTATCTGCTCCAGGCG 1851
QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerThrHisGlyLysSerProPheGly 620
Db 1852 CAGAACCAAGTACAGAGCTGACTCGCGCGAGCTGGCATGAAGAGAGCCCTTTGAA 1911
QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyLysSerHisLeuSerGluAsn 640
Db 1912 AAGCAGTTTAAACGCAAGAGCTGCCAAATGGAATTTGAGAGAGCATCTATGTGAGAGAAC 1971
QY 641 ArgSerArgGlyLysLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db 1972 AGGTCAAGGAGAGCTGGGGAAGTGGGAGTCACTTACTTCTTCTGGGACAGATGGA 2031
QY 661 HisLeuGlyValSer 665
Db 2032 ATCATTTAGAGTCTCC 2046

RESULT 14
US-10-072-012-255
; Sequence 255, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tcheurev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patlurajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimil Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Rutak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsdorok II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Riegey, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

US-10-029-345A-109 (1-665) X US-10-072-012-255 (1-2200)

QY	153	ProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuYrYLeuGlyCysGlnArg	172
Db	541	CTGTTGCGCAACATTGGGCAACCCGAATTCTCCCAATCTTTATCTTGGCTGCCACGA	600
QY	173	AspValLeuAsnIleGlyLeuLeuIleGlnGlnAsnGlyIleGlyYrValLeuAsnAlaSer	192
Db	601	GATGTCTCTCAACAAAGAGCTGATTCGACGAAATGGATATGGTATATGGTAAATGCCACG	660
QY	193	TyrThrCysProIysProAspPheIleProGluSerHisAspLeuArgValProValaen	212
Db	661	AATACCTGTCGAAGCTGACTTATATCCCGAGCTCATCTCTCGGTGGCTGGAT	720
QY	213	AspSerPheCysGlyValIleLeuProThrLeuAspIlySerValAspPheIle-----	230
Db	721	GACAGCTTTGGAGAAATTTTCCGTGGTGGCAAAATCAATCAATGATTTCAATGCTAAG	780
QY	231	-----GluYrValaIysAlaSerAsnGlyCysValLeuValHisCysLeuAla	246
Db	781	TTGACTTATACAGAAAGCAAAAGCCCTCAATGATGTTCTTGTCGTGACATGTTACT	840
QY	247	GlyIleSerArgSerAlaThrIleAlaIleAlaYrIleMetYrArgMetAspMetSer	266
Db	841	GGAATCTCCCTCCCTCCGACCATCGCTATCGCTCATCATGAAAGATGACATGCT	900
QY	267	LeuAspGlnAlaYr---ArgPheValIysGlyValArgProCnrlIleSerProAsnPhe	285
Db	901	TTAATGATGAAGCTTACAGAGATTTTGGAAAGCAAAAGCTTACTATATCTCCAAATCTC	960
QY	286	AsnPheLeuGlyGlnIleLeuAspYrGlyValYrIleYrAsnGlnThrGlyAlaSer	305
Db	961	AATTTTCTGGGCAACCTCTTGACTATAGAAAGAAATTAGAACCAAGACTGGAAGATCA	1020
QY	306	GlyProIysSerIysLeuValIleuGlnIleuGlyValProAsnGluProValProAla	325
Db	1021	GGGCGCAAAAGCAAACTCAAGCTGTGCACTCGAAGAACCAATGAACCTGTCCTGCT	1080
QY	326	ValSerGlyGlyGlyGlyLysSerGluThrProLeuSerProProCysAlaAspSerAla	345
Db	1081	GTCCTGAGGCGTGGACAAAGAAAGGACGCGCTTCAGTCCACCTGTGCGCATCTGCT	1140
QY	346	ThrSerGlnAlaIleGlyGlnArgProValHisProAlaSerValProSerAlaProSer	365
Db	1141	ACCTGAGGGCGACGACCAAAAGCCCGTGCATCCGCGCAGCTGCCAGCTGCCACGC	1200
QY	366	ValGlnProSerLeuLeuGlyAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeu	385
Db	1201	GTGACCGCTGCGCTTGAAGACAGCCGCTGGTACAGGCGCTCAGTGGCTGCACCTG	1260
QY	386	SerIlaAspAlaGlyLeuGluAspSerAsnIlyLeuYrArgSerPheSerLeuAspIleYr	405
Db	1261	TCCGAGACAGGCTGGAGAACACCAATTAAGCTCAACGCGTCTCTCTCGAATACAA	1320
QY	406	SerValSerYrSerAlaSerMetAlaIleSerLeuHisGlyPheSerSerSerGlnAsp	425
Db	1321	TCAGTTTCATATTCAGCGACGACATGGAGATCTTAACTAGGCTTCTCTCATGAAAGAT	1380
QY	426	AlaLeuGluYrYrYrYrYrProSerThrThrIleuAspGlyThrIleAsnIlyLeuGlyAsn	445
Db	1381	GCTTGGAAATACTAACAACTTCTCACTACTCGATGGAGCAACAAAGATATGCAAGTTC	1440
QY	446	SerProValGlnGluLeuSerGluGlnThrProGluThrSerProAspIlySerGluGlnAla	465
Db	1441	TCCCTGTTCCAGAACTATCGAGCAGCTCCGAAACCAAGTCTGATTAAGGAGGAAGCC	1500
QY	466	SerIleProIysLeuLeuGlnThrAlaArgProSerAspSerGlnSerIlyrArgLeuHis	485
Db	1501	AGATATCCCAAAAGACTGAGACCGCCAGAGGCTTTCACACAGCCAGCAAGCGATTGCAT	1560
QY	486	SerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisArg	505
Db	1561	TGAGTCAGAACAGACAGATGGCAGCCGCCAAGAGTCCCTTTATCTCACTGATGCA	1620

Qy	506	SerGIysSerValIGluAspAsnTrpHisSerPheLeuPheGlyLeuSerThrSerGln	525
Db	1621	AGTGGAGACCGTGGAGGACAAATTATTCACACCAAGCTTCCTTTTCGGCTTTCCACACGACG	1680
Qy	526	GlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTrpHisSerAspIleLeuAla	545
Db	1681	CAGCACCTCACGAGACTCTGCTGGCTGGGCTTTAAGGCTGGCACTGGAGATCTTGGCC	1740
Qy	546	ProGlnThrSerThrProSerLeuThrSerSerTrpPheAlaThrGluSerSerHis	565
Db	1741	CCCCAGACCTCTACCCCTCTCCCTGACACGACGCTGGTATTTTGGCACAGAGCTCCACAC	1800
Qy	566	PheTrpSerAlaSerAlaIleTrpGlyGlySerAlaSerTrpSerAlaTySerCysSer	585
Db	1801	TTTCACTTGCTCAAGCCACTTACGAGAGCAGCTGCCACTTACTCTGCTTACAGCTGCACG	1860
Qy	586	GlnLeuProThrCysGlyAspGlnValTrpSerValaArgArgGlnLysProSerAsp	605
Db	1861	CAGCTGCCCATTTGGGGAGACCAAGTCTATTCTGAGCCAGCGGCAGAACCCAAAGTAC	1920
Qy	606	ArgAlaAspSerArgArgSerTrpHisGlnGluLysProPheGlnLysGlnPheLysArg	625
Db	1921	AGACTGTACTGCGCGGAGCGTGGCATTAAGAGAGCCCCCTTTGAAMACAGATTAAAGC	1980
Qy	626	ArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsnArgSerArgGlnGlu	645
Db	1981	AGAAAGCTCCAAATGGAAATTTGGAGAGAGCATCATGTACAGAAACAGGTCACGGGAAAG	2040
Qy	646	LeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGluIleGlnLysSer	665
Db	2041	CTGGGGAAGTGGGACAGTCAAGTCTTAGCTTTTTCGGCACAGTGAATATCTGAGGTCTCC	2100

```

RESULT 15
US-09-964-277-20
; Sequence 20, Application US/09964277
; Patent No. US2002013710A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-20

Alignment Scores:
Pred. No.:          8,86e-317          Length:      3332
Score:              3048.50             Matches:      607
Percent Similarity: 91.44%             Conservative: 2
Best Local Similarity: 91.14%           Mismatches:  1
Query Match:        89.19%             Indels:      56
DB:                  9                  Gaps:         1
US-10-029-345A-109 (1-665) x US-09-964-277-20 (1-3332)

```

Oy	1	MetAlaHisgIumettIeGlyThrGlnIleValThrIuAryGLeuValAlaLeuIeugIu	20
Db	562	ATGGCCCATCGAGATGATTGGAACTCAATTGTACTGAGAGGTGGTGGCTTGGTGGAA	621
Oy	21	SeerGlyThrGluIyValIeuleuIleAspSerAryProPheValGluTyraAsnThrSer	40
Db	622	AGTGGAAACGAAAAAGTCTGCTGATTAATGATGCGGCATTTGGTGAATACAAATCAATCC	681
Oy	41	HisIleIeuGluAlaIleAsnIleAsnCySerIyLeuMetLybAryGLeuIeugIuGln	60
Db	682	CACATTTTGGAGCATTAATATCAACTGCTCCAAAGCTTAATGAACGAGGTGGCAACG	741

QY	61	AspIysValIleuLeuIleThnGluLeuIleGlnHisSerAlaIysHisIleValAspIleLeu	80
Db	742	GACAAAGGTTAAATTACAGACTCATTCAGACATTCAGGAAACATTAAGGTTGACATTGAT	801
QY	81	CysSerGlnIysValIleValIleTyrAspIleSerSerGlnAspValAlaSerIleSerSer	100
Db	802	TGCAGTCAGAAAGGTTGTAGTTTACGATCAAAAGCTCCCAAGATGTTGCCCTCTCTCTCA	861
QY	101	AspCysPheLeuThrValIleuLeuGlyIleValLeuGluIysSerPheAsnSerValHisIleu	120
Db	862	GACGTGTTTTCACGTACTCTTGAGGTAACTGGAGAAAGACTTCAACTGTCTTACCTGG	921
QY	121	-LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGluIleuCysGluGlyIleVal	140
Db	922	CTTG-----	925
QY	140	rThrIleuValProThrCysIleSerGlnProCysIleuProValAlaAsnIleGlyProTh	160
Db	925	-----	925
QY	160	rArgIleuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGluLeuIle	180
Db	926	-----CAGAGACTGAT	936
QY	180	eGlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerTyrThrCysProIysProAspPh	200
Db	937	:GACAGAGATGGGATTGGTTATGTGTTAAAGCCAGCAATCTGTCCAAAGCCCTGACTT	996
QY	200	eIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIleIle	220
Db	997	TATCCCCAGTCTCAATTCCTGCGGTGCGCTGTGATACAGCTTTGTGAGAAATTTT	1055
QY	220	uProThrLeuAspIysSerValAspPheIleGluIleValAlaIysAlaSerAsnGlyCysVa	240
Db	1057	GCCGCGTTGAGCAAAATCAGTAGATTCATTGAGAAACAAAGCCTCCATGATGTGT	1114
QY	240	IleuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleLeu	260
Db	1117	TCTAGTGACATGTTAGCTGGGAGTCTCCCGCTCCGACCAATCGCTACCGCTCATCAT	1176
QY	260	tlValArgMetAspMetSerIleuAspGluAlaTyrArgPheValIleValIysValArgProTh	280
Db	1177	GAAAGGATGAGACATGTCATTAGATGAAGCTTACGATTTGTGAAGAAAGAAAGCTTAC	1233
QY	280	rIleSerProAsnPheAsnPheLeuGlyGlnIleuLeuAspTyrGluIleValIleValAs	300
Db	1237	TATATCTCCAAACTTCATTTTCTGGGCAACTCCTGGACATATGAGAAAGATTAAAGA	1296
QY	300	nGlnThrGlyAlaSerGlyProIysSerIleValLeuIleValLeuGluIleValProAs	320
Db	1297	CCAGACTGAGACATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGGAGAACCAA	1356
QY	320	nGluProValProAlaValIleSerGluGlyGlyGlnIleValSerGluThrProIleSerPro	340
Db	1357	TGAACCTGTCCCTCTGCTCAGAGGGGTGACAGAAAGCGAGAGGCCCTCTCAGTCCACC	1411
QY	340	cCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal	360
Db	1417	CTGTGCGCACTCTGCTACTCAGAGCGACGACGACAAAGGCCCGTGCATCCCGCCAGGCT	1477
QY	360	IProSerValProSerValGlnProSerIleuLeuGluAspSerProIleuValGlnAlaLe	380
Db	1477	GCCACAGCGTCCAGCGTGCAGCGTCCGTGTAGAGGACAGCCGCTGTATCAGGCGCT	1536
QY	380	uSerGlyIleuHisIleuSerAlaAspArgLeuGluAspSerAsnIleValIleValSerSer	400
Db	1537	CAGTGGGCTGACCTGTCCGACAGAGCTGAAAGACGCAATTAAGCTTCAAGCTTCTT	1596
QY	400	eSerIleuAspIleIleValSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPh	420
Db	1597	CTCTCTGATATCAAAATCAATGTTTCAATATTCAGCCAGACATGGCAGCATCTTACATGGCTT	1655
QY	420	eSerSerSerGluAspAlaLeuGluIleTyrIleValProSerThrThrIleuAspGlyThrAs	440

Db 1657 CTCCTATCAGAAAGATGCTTTGGAAATACAAACCTTCCACTACTGATGGAGCCAA 1716
Qy 440 nlyslleucyvglnpheserprovalglnluleusergluglnthrProglumthserPr 460
Db 1717 CAAGCTAATGCAAGTTCCTCCCTGTTCAGGAACATCGAGAGAGATCCCGAACCAGTCC 1776
Qy 460 oAepLySGlUGlAlSerLleProLySLyLeuGlnThrAlaArgProSerAspSerGI 480
Db 1777 TGATTAAGAGGAAGCCAGCATCCCAAGAGAGCTGCAGACCCGACGCTTCAGACAGCA 1836
Qy 480 nSerLySAglleuHlASerValArgThrSerSerSerGIYThralaglnArgSerleuLe 500
Db 1837 GAGCAGAGCATTCGATTCGATCAGAACAGCAGAGAGAGTGGACCCGCCAGAGTCCCTTTT 1896
Qy 500 uSerProLeuHlASrSerGIYSerValGluAspAsnTYrHlsthSerPheLeuPheGI 520
Db 1897 ATCTCCACTGCATCCAAAGTGGAGGCGTGGAGGACAAATTAACAACAGCTTCCCTTTCCG 1956
Qy 520 yLeuSerThrSerGIInHlslLeuThrLySerAlaglyLeuGlyLeuLySGlyTrpHl 540
Db 1957 CTTTCCACAGCCAGCCAGCAGCACTCAAGAGTCTGCTGGCTGGCTTAAAGGCTGGCA 2016
Qy 540 sSerAspLleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTYrPheAl 560
Db 2017 CTGGGATATCTTGGCCCCCAAGACTCTACCTTCCCTGACACAGCAGCTGGTATTTTGC 2076
Qy 560 aThrGIuSerSerHlAsPheTYrSerAlaseralalleTYrGlyGIYSerAlasertYrSe 580
Db 2077 CACAGAGTCCCAACACTTACTCTGCTCAGCCATCTAAGAGAGCAGTGCAGTTACTTC 2136
Qy 580 rAlaTYrSerCyserGIInleuProthryrsglyAspGlnValTYrSerValArgArgAr 600
Db 2137 TGCCTACAGCTGCAGCCAGCTGCCCACTTGCAGAGACCAAGTCTATTCTGTGCGCAGCG 2196
Qy 600 gglnlySProSerAspArgAlaAspSerArgArgSerTrpHlsglUGluserProPheGI 620
Db 2197 GCAGAAAGCCAAAGTGAAGAGTGAATCGCGGAGCTGGCANTGAAGAGAGCCCTTTGA 2256
Qy 620 ulYsGlnPheLyArgArgSerCySglnMetGlnPheGIYGIuserllemetSerGIuAs 640
Db 2257 AAAGCAGTTTAAACGACGAAGCTGCCAATGGAATTTGGAGAGAGCATGATCAGAGAA 2316
Qy 640 nArgSerArgGlnGlnluleuGlyLyValGlySerGlnSerSerPheSerGIYserMetGI 660
Db 2317 CAGGTACCGGAAAGAGCTGGGAAAGTGGGACAGTCAAGTCTTTCGGGACAGCATGA 2376
Qy 660 uileileGIuValser 665
Db 2377 AATCATGAGTCTCC 2392

Search completed: June 22, 2004, 03:06:35
Job time : 857.731 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 19:56:46 ; Search time 735.833 Seconds
(without alignments)
3839.259 Million cell updates/sec

Title: US-10-029-345A-109
Perfect score: 3418
Sequence: 1 MAHEMCTQVTERVALLE.....LGKVGSGSRFGSGMEITIEVS 665

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xjh
-Q/cgnt_1/USPTO.spool/US10029345/runat_21062004_164145_29020/app_query.fasta_1.1294
-DB=Geneseq_29Jan04 -OPMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345 @CGN_1_1_687 @runat_21062004_164145_29020 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2003as: *
8: geneseqn2003bs: *
9: geneseqn2003cs: *
10: geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3418	100.0	5450	6	ACC60559 Polynucle
2	3418	100.0	5450	6	ACC60572 Polynucle
3	3406	99.6	2732	4	AAD09492 Human SGP
4	3406	99.6	3059	6	AAS15768 CDNA seqn
5	3406	99.6	3104	6	ABN59704 Novel hum
6	3406	99.6	3496	6	ABK47596 CDNA enco
7	3406	99.6	3544	5	AAS14639 Human CDN
8	3406	99.6	3544	6	ABK49402 CDNA enco

153-277 Check

9	3406	99.6	3766	6	ABK14474	ABK14474 Human pro
10	3406	99.6	4790	6	ABN83966	ABN83966 Human gen
11	3406	99.6	5145	5	ABV20833	ABV20833 Human pro
12	3406	99.6	5145	5	ABV21080	ABV21080 Human pro
13	3406	99.6	5145	5	ABV26680	ABV26680 Human pro
14	3406	99.6	5145	5	ABV20978	ABV20978 Human pro
15	3406	99.6	5145	5	ABV21092	ABV21092 Human pro
16	3406	99.6	5145	5	ABV21312	ABV21312 Human pro
17	3406	99.6	5145	5	ABV21316	ABV21316 Human pro
18	3406	99.6	5145	5	ABV26826	ABV26826 Human pro
19	3406	99.6	5145	5	ABV27131	ABV27131 Human pro
20	3406	99.6	5145	5	ABV26923	ABV26923 Human pro
21	3406	99.6	5145	5	ABV27135	ABV27135 Human pro
22	3406	99.6	5145	5	ABV28657	ABV28657 Human pro
23	3406	99.6	5145	5	ABV22827	ABV22827 Human pro
24	3406	99.6	5145	5	ABV26934	ABV26934 Human pro
25	3399	99.4	2102	7	ADA53105	ADA53105 Human cod
26	3398	99.4	2966	4	AAH99685	AAH99685 Human pro
27	3388	99.1	5111	6	ACC60521	ACC60521 Polynucle
28	3373	98.7	2118	4	AAF30479	AAF30479 Human pro
29	3057.5	89.5	2756	6	ACC60560	ACC60560 Polynucle
30	3048.5	89.2	3332	6	ABK48378	ABK48378 CDNA enco
31	1326	38.8	2377	7	ACA64956	ACA64956 Human pro
32	1326	38.8	2377	7	ABX10760	ABX10760 Human pro
33	1302	38.1	2453	6	AAD41236	AAD41236 Murine ne
34	1297	37.9	2453	2	AAT66758	AAT66758 Human pro
35	1262	36.9	2476	4	AAS31013	AAS31013 Human dia
36	1246	36.5	749	4	AAH06539	AAH06539 Human cod
37	1164.5	34.1	2415	2	AAT676757	AAT676757 CDNA of t
38	977.5	28.6	1917	9	ADDE07547	ADDE07547 Novel cod
39	742.5	21.7	1171	5	AAS72089	AAS72089 DNA encod
40	724	21.2	418	5	ABV11252	ABV11252 Human pro
41	724	21.2	461	5	ABV40981	ABV40981 Human pro
42	724	21.2	461	5	ABV42398	ABV42398 Human pro
43	724	21.2	461	5	ABV41327	ABV41327 Human pro
44	707	20.7	425	5	ABV10726	ABV10726 Human pro
45	703	20.6	467	5	ABV40998	ABV40998 Human pro

ALIGNMENTS

RESULT 1	ACC60559	standard; CDNA; 5450 BP.
ID	ACC60559	
AC	ACC60559;	
XX		
DT	19-JUN-2003	(first entry)
XX		
DE	Polynucleotide relating to the invention SEQ ID NO: 108.	
XX		
KW	Gene; ss; antiproliferative; hepatotropic; nephrotropic; antirheumatic;	
KW	antiproliferative; cardiatic; cytosolic; gene therapy; liver disease;	
KW	proliferative disorder; renal failure; cardiovascular disorder;	
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;	
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200257460-A2.	
XX		
PD	25-JUL-2002.	
XX		
PF	20-DEC-2001; 2001WO-US050459.	
XX		
PR	20-DEC-2000; 2000US-0256868P.	
PR	30-MAR-2001; 2001US-0280186P.	
PR	01-MAY-2001; 2001US-0280773P.	
PR	05-JUN-2001; 2001US-0295848P.	
PR	25-JUN-2001; 2001US-0300465P.	
XX		
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.	
XX		

Db	2098	CTTTCACACGACGACGACGACCTACGAAAGCTGCTGGCTCGGGGCTTAAAGGCTGGAC	2157
Qy	541	SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla	560
Db	2158	TCGGATATCTTGGGCCCCCAGACCTCTACCCCTTCCCTGACACAGACGCTGGATTGTC	2217
Qy	561	ThrGluSerSerHisPheTyrSerSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer	580
Db	2218	ACAGAGTCTTACACTTACTCTGCTGCTCACGACCACTAGGAGGAGGAGCCAGTTACTCT	2277
Qy	581	AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg	600
Db	2278	GCCATACAGCTGCACACGACGCTCCCACTTGGCGAAGCAAGCTATTTCTGTGCGCAGGCG	2337
Qy	601	GlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGluGluSerProPheGln	620
Db	2338	CAGAAAGCCAAATGACAGAGCTGACTCGCGGCGAGAGCTGCATGAAGAGAGCCCTTTGAA	2397
Qy	621	LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn	640
Db	2398	AAGAGATTAAACCGACGAAAGCTGCCAAATGGAAATTTGGAGAGACATCATGTACAGAGAC	2457
Qy	641	ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu	660
Db	2458	AGGTACACGGAGAGCTGGGGAAAGTGGGACAGTCACTTACGCTTTTCGGGACAGATGAA	2517
Qy	661	IleIleGluValSer	665
Db	2518	ATCATTTAGGCTCTCC	2532
RESULT 2			
ACC60572	ACC60572	standard; cDNA; 5450 BP.	
AC	ACC60572;		
XX	19-JUN-2003	(first entry)	
DE	Polynucleotide relating to the invention SEQ ID NO: 147.		
XX	Gene; 68; antiproliferative; hepatotropic; nephrotoxic; antiarthritic;		
KM	antipneumatic; cardiant; cytostatic; gene therapy; liver disease;		
KM	proliferative disorder; renal failure; cardiovascular disorder;		
KM	immunological disorder; arthritis; psoriasis; congenital heart defect;		
XX	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.		
OS	Homo sapiens.		
XX	MO200257460-A2.		
XX	25-JUL-2002.		
PF	20-DEC-2001;	2001WO-US050459.	
XX	20-DEC-2000;	2000US-025686P.	
PR	30-MAR-2001;	2001US-0280186P.	
PR	01-MAY-2001;	2001US-0287735P.	
PR	05-JUN-2001;	2001US-0295848P.	
PR	25-JUN-2001;	2001US-0300465P.	
XX	(BRIM) BRISTOL-MYERS SQUIBB CO.		
XX	Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;		
PI	Siemens N, Bol D, Schlieven G, Finger J, Todderud CG, Bassolino D;		
PI	Krystek S, Mcatee P, Suchard S, Banas D;		
XX	WPI; 2002-599721/64.		
DR	P-PSDB; ABR52407.		
PT	Novel polynucleotides encoding human phosphatase polypeptides useful in		
PT	the prevention or treatment of e.g. proliferative and cardiovascular		
XX	disorders.		

Example 7: Fig 19; 801bp; English.

xx The invention relates to a novel isolated nucleic acid comprising a
cc polynucleotide having a nucleotide sequence selected from 40
cc polynucleotides fully defined in the specification. The polynucleotide of
cc the invention has antiproliferative, hepatotropic, nephrotropic,
cc antiarthritic, antipsoiastic, cardiant, and cytostatic activity. The
cc polynucleotide may have a use in gene therapy. A polynucleotide or
cc polypeptide of the invention is useful for preventing, treating or
cc ameliorating a medical condition, e.g. a proliferative disorder. They are
cc also useful for treating e.g. liver disease, renal failure, immunological
cc disorders including arthritis and psoriasis, cardiovascular disorders
cc such as congenital heart defects and congestive heart failure, and
cc cancer. A method of the invention is useful for diagnosing a pathological
cc condition or susceptibility to a condition in a subject. The present
cc sequence is used in the exemplification of the invention

xx
sq Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.02e-235	Length:	5450
Score:	3418.00	Matches:	665
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345A-109 (1-665) x ACC60572 (1-5450)

QY 1 MetAlaHisgIuMeCllEgIYThrGlnIleValThrGluArgLeuValAlaIleuLeuGlu 20

Db 538 ATGGCCCATGAGATGATTGGAACTCAATATTGTACTGAGAGCTTGCTGCTGCTGGAA 597

QY 21 SerGlyThrGluYsValLeuLeuIleAspSerArgProPheValGluTYrAsnThSer 40

Db 598 AGTGAACCGGAAAAAGCTGCTCTAATGATGCGCGCATTTGGGATATACAAATACATCC 657

QY 41 HisIleuGlnAlaIleAsnIleAsnCySerYsIleuMeClyAsArgLeuGlnIn 60

Db 658 CACATTTTGGAGCCATTAAATATCACTGCTCCAACTTATGAAAGCGAGTTGCCAAG 717

QY 61 AspIysValLeuIleThrGluLeuIleGlnHisSerAlaYsHisIysValAspIleAsp 80

Db 718 GACAAAGGTTAATTACAGAGCATCATCAGACATTACAGGAAACATTAAGTTGACATTGAT 777

QY 81 CysSerGlnYsValValValTYrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100

Db 778 TGCAGTCAAGAGGTGTGATGTACAGATCAAGCTCCCAAGATGTGCTCTCTCTTCA 837

QY 101 AspCyPheLeuThrValLeuLeuGlyLYsLeuGlnYsSerPheAsnSerValHisIleu 120

Db 838 GACTGTTTTCTCATCTGACTTCTTGGGTTAACTGGAGAAAGACTTCAACTCTGTTCACTG 897

QY 121 LeuAlaGlyIYpheaIaGluPheSerArgCysPheProGlyLeuCyGluGlyLYsSer 140

Db 898 CTTGAGAGTGCTTTGCTGAGATTCTCTCGTTGTTCTCTGCTGCTCTGGAAGAAATCC 957

QY 141 ThrIleuValProThrCySllSerGlnProCysIleuProValAlaAsnIleGlyProThr 160

Db 958 ACTGTAGCTCCATCTGATTTCTCAGCTTGCTGATCTGTGTCCAAATTTGGGCAACC 1017

QY 161 ArgIleuProAsnLeuTYrLeuGlyCYsGlnArgAspValIleAsnIysGluIleu 180

Db 1018 CGAATTTCTCCCATCTTTATCTTGCGTCCAGAGAGATGTCCTCAACAAGAGACTGATA 1077

QY 181 GlnGlnAsnGlyIleGlyTYrValLeuAsnAlaSerTYrThrCySProlYsProAspPhe 200

Db 1078 CAGCAGATGGAGATTGGTTATGTATTAATGCCAGTATACCTGTCCAAACCTGCATCTT 1137

QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCySgluYsIleu 220

Db 1138 ATCCCGAGTCTCATTTCTCGCGGTGCTCTGAAATGACACTTTTGTGAAGAAATTTTG 1197

QY 221 ProTribLeuAspLysSerValaAspPhe11eGluLysAlaLysAsnGlyCysVal 240
 Db 1198 CCGTGGTTGGACAAATAGTAGATTTCATTGAGCAAAAGAACCTCCAAATGATGCTTT 1257
 QY 241 LeuValH1sCysLeuAlaGly11eSerArgSerAlaThr11eAla11eAla11eMet 260
 Db 1258 CTAGTGACATGTTTACCTGGGATCTCCCGCTCCGACCATGCTATGCGCTACATCATG 1317
 QY 261 LysArgMetAspMetSerLeuAspGlu1a1YrArgPheValLysGluLysArgProThr 280
 Db 1318 AAGAGATGACATGCTCTTAAATAGAGCTTACATTTGTAAGAAAGAAAGAAAGCTACT 1377
 QY 281 11eSerProAsnPhenAsnPhenLeuGly11eLeuAsp1YrGluLys11eLysAsn 300
 Db 1378 AATATCTCCAAACTTCAATTTCTGGGCCAATCTCGAGCTATGAGAAAGAAATTAAGAAC 1437
 QY 301 GlnThrGlyAlaSerGly1ProLysSer1LysLeuLysLeuH1sLeuGluLysProAsn 320
 Db 1438 CAGATCGAGCATTCAGGGCCAAAGCAAACTCAAGCTGCTGCACCTGGAGAACCAAT 1497
 QY 321 GluProVal1ProAlaValaSerGluGly1Gly1LysSerGluThrProLeuSerProPro 340
 Db 1498 GAACCTGTCCTGCTGCTCAGAGGGTGGACAGAAAGCAAGCCCTCACTCCACCC 1557
 QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGly1Gly1ArgProValH1sProAlaSerVal 360
 Db 1558 TGTGCCGACTCTGCTCACTCAGAGGACAGAGCAAAAGGCCGTGCATCCCGCAGCGTG 1617
 QY 361 ProSerVal1ProSerVal1GlnProSerLeuLeuGluAspSerProLeuVal1GlnAlaLeu 380
 Db 1618 CCCAGCGTGCCAGCGTGACGCGTGCTTTAGAGAGACGCCGCTGTATACAGCGCTC 1677
 QY 381 SerGlyLeuH1sLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
 Db 1678 AGTGGGCTGACCTGTCCGACAGCGCTGAGAACAGACAAATAGCTCAAGGTTCTTC 1737
 QY 401 SerLeuAsp11eLysSerValSer1YrSerAlaSerMetAla1a1eSerLeuH1sGlyPhe 420
 Db 1738 TCTCTGATATCAAAATCAGTTTCATATTCAGCCACATGCGACATCTTTCATGAGCTTC 1797
 QY 421 SerSerSerGluAspAlaLeuGlu1Yr1YrLysProSerThrThrLeuAspGly1ThrAsn 440
 Db 1798 TCTCTATCAAGAAATGCTTGGAAATACAAACCTTCCACTCTGAGATGGAGCCAC 1857
 QY 441 LysLeuCysGlnPheSerProVal1GlnLysLeuSerGluGlnThrProGluThrSerPro 460
 Db 1858 AAGCTATGCGAGTTCTCCCTGTTTCAAGAACTATTCAGAGACTCCCGAAACCAAGTCTT 1917
 QY 461 AspLysGluGluAlaSer11eProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 Db 1918 GATTAGAGAGAAAGCCAGATCCCAAGAGCTGACAGCCCGCAGGCTTCAGACACCCAG 1977
 QY 481 SerLysArgLeuH1sSerValArgThrSerSerSerGly1YrAlaGlnArgSerLeuLeu 500
 Db 1978 AGCAAGCATTTGCAATTCGTTCAGAACCAAGCAAGCAAGTCCCGCAGAGTCCCTTTTA 2037
 QY 501 SerProLeuH1sArgSerGlySerVal1GluAspAsn1YrH1sThrSerPheLeuPheGly 520
 Db 2038 TCTCACTGCAATCGAAATGGAGGTGGAGAGCAATTAACACCAAGCTTCTTTTCCGC 2097
 QY 521 LeuSerThrSerGlnGlnH1sLeuThrLysSerAlaGly1LeuGly1LeuLysGly1YrH1s 540
 Db 2098 CTTTCCACACCGACGACGACCTCAAGAGTCTGCTGCGCTTAAGAGGCTGGGAC 2157
 QY 541 SerAsp11eLeuAlaProGlnThrSerThrProSerLeuThrSerSer1YrPheAla 560
 Db 2158 TCGGATATCTTGGCCCCCAGACCTTACCTCCCTCCGACACAGAGCTGTGATTTTCC 2217
 QY 561 ThrGluSerSer1sPhe1YrSerAlaSerAla1e1YrGly1YrSerAlaSer1YrSer 580
 Db 2218 ACAGAGTCTCCACACTTACTCTGCTCAGCACTTACGAGAGCGAGCGACAGTACTCT 2277
 QY 581 Ala1YrSerCysSerGlnLeuProThrCysGlyAspGlnVal1YrSerValArgArgArg 600

Db 2278 GCCTACAGCTGACGACGACTGCCACTTCCGAGAACCAAGCTATTTCTGTGCGCAGCGG 2337
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTrpH1sGluGluSerProPheGlu 620
 Db 2338 CAGAAAGCAAGTGAAGAGCTGACTCGCGCGAGCTGGCATGAGAGAGCCCTTTGA 2397
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGly1GluSer11eMetSerGluAsn 640
 Db 2398 AAGCATTTAAACCAAGAGCTGCCAAATGAAATTTGAGAGAGCATATGTCAGAGAAC 2457
 QY 641 ArgSerArgGluGluLysValaGlySerGlnSerSerPheSerGlySerMetGlu 660
 Db 2458 AGGTCAAGGAGAGCTGGGAGAAAGTGGAGTCAAGTCACTTCTTCCGAGCATGGA 2517
 QY 661 11e1eGluValSer 665
 Db 2518 ATCATTTGAGGCTTCC 2532
 RESULT 3
 AAD09492
 ID AAD09492 standard; DNA; 2732 BP.
 XX AAD09492;
 AC
 AC
 AC
 DT 10-SEP-2001 (first entry)
 DE
 DE
 XX Human SGP002 phosphatase polypeptide encoding DNA.
 XX
 KW Human; SGP002 phosphatase polypeptide; phosphatase-related disease;
 KW immune-related disorder; ocular disease; organ transplant rejection;
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
 KW metabolic disorder; haematopoietic cancer; mood disorder; cardiact;
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
 KW attention disorder; cognition disorder; psychotic disorder; cytoskeletal;
 KW neurological disorder; viticidal; nocturnal; cerebroprotective; therapy;
 KW neuroprotective; antibacterial; vulnary; transilliser; anasthetic;
 KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
 KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
 KW MKP; mygraine; chromosome 12p11.1-p12.1; ds.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 538..2535
 FT /*tag= a
 FT /product= "Human SGP002 phosphatase polypeptide"
 PN WO200146394-A2.
 XX
 XX
 PD 28-JUN-2001.
 XX
 XX
 PP 21-DEC-2000; 2000MO-US034736.
 XX
 XX
 PR 21-DEC-1999; 99US-017325SP. NO
 PR 28-DEC-1999; 99US-0175766P. NO
 PR 25-JAN-2000; 2000US-0178078P. NO
 PR 31-JAN-2000; 2000US-0179301P. NO
 XX
 XX (STGE-) SUGEN INC.
 PA
 XX
 XX
 PI Pjowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;
 PI Hill RJ, Flanagan P;
 DR MPI; 2001-418058/44.
 DR P-PSDB; AAE04834.
 XX
 XX Novel phosphatase polypeptide useful for treating cancers, immune-related
 PT diseases and disorders, cardiovascular disease, brain or neuronal-
 PT associated diseases and metabolic disorders.
 XX
 XX Claim 29; Fig 1; 166pp; English.

XX The present invention relates to phosphatase polypeptides, nucleotide sequences encoding them, as well as various products and methods useful for the diagnosis and treatment of various phosphatase-related diseases and conditions. Substance that modulates the activity of phosphatase polypeptide is used to treat immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases and metabolic disorders, including cancers of tissues, cancers of haematopoietic origin, diseases of central and peripheral nervous system, Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, viral infections, infections caused by prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual dysfunction, mood disorders, attention disorders, cognitive disorders, hypotension, hypertension, psychotic disorders, neurological disorders, dyskinesias and organ transplant rejection. The present sequence is a DNA encoding human SGP002 phosphatase polypeptide. This sequence is classified as dual specificity phosphatase (DSP) and Map kinase phosphatase (MKP). SGP002 gene maps to chromosomal position 12p11.1-p12.1

XX Sequence 2732 BP, 710 A; 684 C; 686 G; 652 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8,73e-235	Length:	2732
Score:	3406.00	Matches:	663
Percent Similarity:	99.85%	Conservative:	1
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	99.65%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-109 (1-665) x AMD09492 (1-2732)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 DB 538 ATGGCCCATGATGATTTGGAACTCAATGTTAATGAGAGTGGGCTGCTGCGGAA 597
 QY 21 SerG1ThrG1uysValLeuLeuIleAspSerArgProPheValGluTyraAnthSer 40
 DB 598 AGTGGACGAAAGAGCTGCTTATTGATAGCCGCAATTTGGAAATTCATATACATCC 657
 QY 41 HisIleLeuGluAlaIleAsnIleAsnCyseSerIysLeuMetIysArgArgLeuGln 60
 DB 658 CACATTTGGAGCATTAATATCACTGCTCCAAAGTTATGAAGCAAGGTTGCAACAG 717
 QY 61 AspIysValLeuIleThrGluLeuIleGlnHisSerAlaIysHisIysValAspIleAsp 80
 DB 718 GACAAAGTGTAAATTAACAGAGCTATCCACATTCACAGCAATAGGTGACATTTGAT 777
 QY 81 CyseSerGlnIysValValIleTyraAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 778 TGCAGTCAGAAAGTGTGATTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 837
 QY 101 AspCysePheLeuThrValLeuLeuGlyIysLeuGluIysSerPheAsnSerValHisLeu 120
 DB 838 GACTGTTTCTCACTGTAATCTTGGGTAACTGGAGAAAGAGCTTCAACTGTTCACTG 897
 QY 121 LeuAlaGlyIysPheAlaGluPheSerArgCysePheProGlyLeuCyseGluGlyIysSer 140
 DB 898 CTTCAGAGTGGTGTGCTGAGTCTCTCGTGTGTTCCCTGGCCCTGTGAGAGAAATCC 957
 QY 141 ThrLeuValProThrCyseIleSerGlnProCyseLeuProValAlaAsnIleGlyProThr 160
 DB 958 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGGCTTACTTGGCAACATTTGGCCAAC 1017
 QY 161 ArgIleLeuProAsnLeuTyraLeuGlyCyseGlnArgAspValLeuAsnIleGlyLeu 180
 DB 1018 CGAATTCCTCCCATCTTTATCTTGGCTGCGAGGAGATGATCCCAACAGAGAGCTGATG 1077
 QY 181 GlnGlnAsnGlyIleGlyTyraLeuAsnAlaSerTyraThrCyseProIysProAspPhe 200
 DB 1078 CACGAGATGGAGATGTGTATGTGTAATGCCAGCAATACCTGTCCAAAGCCGTGACTTT 1137
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCyseGluIysIleLeu 220

DB 1138 ATCCCGAGTCTCATTTCTGCTGCTGCTGATATGACAGCTTTTGTGAGAAATTTTG 1197
 QY 221 ProTTrpLeuAspIysSerValAspPheIleGluIysAlaIysAlaSerAsnGlyCyseVal 240
 DB 1198 CCTGTGTGGCAAAATCAGTAAATTTTCATTTGAAAGCAAAAGCCCTCAATGATGTGTT 1257
 QY 241 LeuValHisCyseLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyraIleMet 260
 DB 1258 CTATGCTACTGTTTACTGAGATCTCCGCTCCGCCACATCCCTATGCTTACATCATG 1317
 QY 261 LysArgMetAspMetSerLeuAspGluAlaTyraArgPheValIysGluIysArgProThr 280
 DB 1318 AAGAGATGAGATGATCTTATAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 1377
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnIleLeuAspTyraGluIysIleIysAsn 300
 DB 1378 ATATCTCCAAACTTCAATTTTCGGGCGCACTCTGACATATGAGAAAGATTAAGAAC 1437
 QY 301 GlnThrGlyAlaSerGlyProIysSerIysLeuIysLeuLeuHisIleGluIysProAsn 320
 DB 1438 CAGACTGAGACATCAGGCGCCAAAGCAACTAAGCTGCTGCACTGGAGAAAGCCAAAT 1497
 QY 321 GluProValProAlaValSerGluGlyGlnIysSerGluThrProLeuSerProPro 340
 DB 1498 GAACGTGCTCCTGCTGCTCAGAGGGTGGACAGAAAAGCAGACGCCCTCAATCCACCC 1557
 QY 341 CyseAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSerVal 360
 DB 1558 TGTGCCACTGCTGCTACCTCAGAGGAGCAGAGCAAGAGCCGCTGATCCCGCAGGCTG 1617
 QY 361 ProSerValProSerValGlnProSerLeuGluAspSerProLeuValGlnAlaLeu 380
 DB 1618 CCGAGGTGCGCAGCGGTGAGCGCTGCTGTATGAGACAGCCCGCTGTACAGCGCTTC 1677
 QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIysLeuIysArgSerPhe 400
 DB 1678 AGTGGGCTGACCTGTCGCGGAGCAGGCTGGAGAGACAGCAATAGCTCAAGCGTTCTTC 1737
 QY 401 SerLeuAspIleIysSerValSerTyraSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
 DB 1738 TCTCTGATATCAAAATCAGTTTATATTCATTCAGCAGATGACACATCTTACATGCTTC 1797
 QY 421 SerSerSerGluAspAlaLeuGluTyraTyraIysProSerThrThrLeuAspGlyThrAsn 440
 DB 1798 TCTCTATCAGAAATGCTTTGAAATCTACAACTTCCACTGCTGATGGAGACCAAC 1857
 QY 441 LysLeuCyseGlnPheSerProValGlnIleLeuSerGluGlnThrProGluThrSerPro 460
 DB 1858 AAGCTATGCAAGTCTCCCTGTTCAAGAACTATCGAGAGACATCCGAAACCAATCCT 1917
 QY 461 AspIysGluGluAlaSerIleProIysIysLeuGlnThrAlaAspProSerAspSerGln 480
 DB 1918 GATTAAGAGAGAACCCAGCATCCCAAGAGCTCAGACTGCGGCTTACAGACAGCAG 1977
 QY 481 SerIysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 DB 1978 AGCAAGAGATGATTCGTCAGAACCCAGCAGAGTGGCCCGCCAGAGAGTCCCTTTTA 2037
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTyraHisThrSerPheLeuPheGly 520
 DB 2038 TCTCCATGATGATGAGATGGAGAGCGTGAAGACATATTCACACACAGCTTCCTTTTGGC 2097
 QY 521 LeuSerThrSerGlnGlnHisLeuThrIysSerAlaGlyLeuGlyIysGlyIysPheHis 540
 DB 2098 CTTTCCACAGCAGCAGACACTCAGAGTCTGCTGCGCTGAGCTTAAAGGCTGGCAC 2157
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyraIysPheAla 560
 DB 2158 TCGAATATCTGGCCCCCAGACCTTACCTTCCCTGACACAGACAGCTGATTTTGGC 2217
 QY 561 ThrGluSerSerHisPheTyraSerAlaSerAlaIleTyraGlyIysSerAlaSerTyraSer 580
 DB 2218 ACAGAGTCCCAACATTTACTGCTGCTGAGCATTCAGGAGGAGGAGTGCAGTTACTCT 2277

QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
Db 2278 GCTTACAGCTTACAGCTGCTCCACATCTGGAGAGACCAAGTCTATTCTGTGGCAGGCG 2337
QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGlnGluSerProPheGlu 620
Db 2338 CAGAAAGCAAGTGCACAGAGCTGACTGCGCGGAGCTGGCATGAGAGAGAGAGAGAGAGAG 2397
QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
Db 2398 AAGAGATTAAAGCCAGAGAGCTGCAAAATGGAATTTGAGAGAGAGCATGTCAGAGAGAC 2457
QY 641 ArgSerArgGlnGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db 2458 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGCACTGCTAGCTTTTGGGAGCATGAGAA 2517
QY 661 IleIleGluValSer 665
Db 2518 ATCATTGAGGTCTCC 2532

RESULT 4

AA515768
ID AA515768 standard; cDNA; 3059 BP.

XX AA515768;

DT 18-JUN-2002 (first entry)

DE cDNA sequence of human (dual specificity phosphatase) DUSP-10.

XX Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;
XX neuronal degeneration syndrome; Alzheimer's disease; depression;
XX schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
XX osteoporosis; diabetes; gene therapy; chromosome 12; ss.
OS Homo sapiens.

PH Key Location/Qualifiers
FT 127..2124
FT /*tag= a
FT /product= "DUSP-10 protein"

XX WO200177340-A1.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-EP003966.

XX 10-APR-2000; 2000EP-00107143.

XX (MERE) MERCK PATENT GMBH.

XX Duecker K;

XX WPI, 2002-010917/01.

XX P-PSDB; AAU09946.

XX Novel dual specificity phosphatase polypeptides useful for treating
XX cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
XX disease, depression, schizophrenia, asthma and immune disorders.
PS Claim 5; Page 34-37; 43pp; English.

XX The present invention relates to a new isolated dual specificity
XX phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
XX sequence that is fully defined in the specification. The invention also
XX provides a sequence encoded by a 3059 nucleotide sequence fully defined
XX in the specification, and a sequence having at least 95 % identity to the
XX polypeptide, or fragments or variants of DUSP-10. The invention is useful
XX for treating cancer e.g. leukemia, colon carcinoma, lung cancer,
XX prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
XX stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,

CC schizophrenia, cardiac myopathies, asthma, immune disorders,
CC inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
CC osteoporosis, diabetes and diabetes associated diseases. The molecules of
CC the invention are also useful as vaccines for inducing immunological
CC response in a mammal, in disease diagnosis and in assays for screening
CC agonistic or antagonistic compounds. Other uses of the invention include
CC identifying membrane bound or soluble receptors, as a diagnostic reagent,
CC in chromosome localisation studies, and as a valuable tool in tissue
CC expression studies. The present sequence represents cDNA of the human
CC dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome
XX 12

SQ Sequence 3059 BP; 831 A; 761 C; 709 G; 758 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.02e-234 Length: 3059
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: Gaps: 0

US-10-029-345a-109 (1-665) x AA515768 (1-3059)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 127 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 186
QY 21 SerGlyThrGlnLysValLeuLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db 187 AGTGAACGAGAAAG 246
QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGlnGln 60
Db 247 CACATTTTGGAGCCATTAATATATCACTGCTCCAGCTTATGAGAGAGAGAGAGAGAGAG 306
QY 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 307 GACAAAGTTTAT 366
QY 81 CysSerGlnLysValAlaValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 367 TGCAGTCAAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 427 GACTGTTTCTCAGCTGATCTTCTGGGTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
QY 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuCysGlnGlyLysSer 140
Db 487 CTTCAGAGTGGGTTGCTGATGCTCTCGTTGTTCCCTGAGCTCTGAGAGAGAGAGAGAGAGAG 546
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 547 ACTTATGCTCTTACCTGATCTTCTGAGCTTCTTACCTGATCTTCTGAGAGAGAGAGAGAGAG 606
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGlnLeuIle 180
Db 607 CGAATCTTCCCAATCTTATATCTTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
QY 181 GlnGlnAspGlyIleGlyTyrValLeuAsnAlaSerTyrTrpCysProLysProAspPhe 200
Db 667 CAGCAGAAAG 726
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlyLysIleLeu 220
Db 727 ATCCCGAGTCCATTTCTGCGTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 786
QY 221 ProTyrPheAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
Db 787 CCGTGTGAGCAAAATCAGTATGATTTTCAATGAGAAACAAACCTTCATGATGATGATGATG 846
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260


```

Db      847 CTATGACACTGTTAGTGGGATCTCCCGCTCCGCCACCATCGCTACATCATG 906
Qy      261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleuSerGluAsn 280
Db      907 AAGAGATGAGCATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGAACCTACT 966
Qy      281 IleSerProMetPheMetLeuGluIleuLeuMetPheTyrGluValIleuAsn 300
Db      967 ATATCTCCAACTCATTTCTTGCGCAACTCTGAGCATATGAGAAAGATTAAAGAAC 1026
Qy      301 GlnThrGluAlaSerGlyProLysSerLeuIleuLeuMetIleuGluIleuProAsn 320
Db      1027 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAGCTGCTGCACTCGAGAAACCAAT 1086
Qy      321 GluProValProAlaValSerGluGlyIleuLysSerGluThrProLeuSerProPro 340
Db      1087 GAACGTGTCCCTGCTGCTCAGAGGGTGAAGAAAAGCAGAGCGCCCTCAGTCCACC 1146
Qy      341 CysAlaAspSerAlaThrSerGluAlaAlaGlyIleuArgProValHisProAlaSerVal 360
Db      1147 TGTGCCACTCTCTACTCAGAGGCGAGAGCAAAAGCCCGCTGCACTCCGCGAGCGTG 1206
Qy      361 ProSerValProSerValGlnProSerLeuGluLysAspSerProLeuValGlnAlaLeu 380
Db      1207 CCGAGCTGCGCCAGCGCGAGCGCGCTGCTTAGAGACAGCCGCTGTACAGCGCGCTC 1266
Qy      381 SerGlyLeuHisLeuSerAlaAspArgLeuLysAspSerAlaValSerPhe 400
Db      1267 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATAAAGCTCAAGCGTTCCTTC 1326
Qy      401 SerLeuAspIleuLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
Db      1327 TCTCTGATATCAATCATGTTTCAATATTCAGCAGCATGCGACATCTTCAACATGCGCTTC 1386
Qy      421 SerSerSerGluAspAlaLeuGluTyrTyrTyrPheProSerThrThrLeuAspGlyThrAsn 440
Db      1387 TCCTCATCAAGAAATGCTTTGGAACTACTCAAAACCTTCCACTGCTGATGGAGCAAC 1446
Qy      441 LysLeuGlnPheSerProValGlnIleuSerGluIleuThrProGluThrSerPro 460
Db      1447 AAGCTATGCGAGTTCTCCCTGTTCAAGAACTATCGGAGAGACTCCCGAAACAGAGCTC 1506
Qy      461 AspGlyGluIleuAspIleuProLysLeuGlnThrAlaArgProSerAspSerGln 480
Db      1507 GATAGAGAGAACCCAGCATCCCAAGAACCTGAGAGCCGCGAGCTTCAAGACAGCAG 1566
Qy      481 SerIleuArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
Db      1567 AGCAAGCGATTGCAATTCGTCAAGAACAGAGAGTGGACCGGCCAGAGGTCCTTTTA 1626
Qy      501 SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
Db      1627 TCTCCACTGATCGAAGTGGAGGTGGAGAGCAATTAACACACAGCTTCTTTTGGC 1686
Qy      521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrHis 540
Db      1687 CTTTCCACCGACGACGACCTTCAACAAATCTGCTGGCCCTTAAAGGCTGGCAC 1746
Qy      541 SerAspIleuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
Db      1747 TCGGATATCTGGGCCCCCGACACTTCAACCCCTTCCCTGACAGAGAGCTGATTTTGGC 1806
Qy      561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAlaSerTyrSer 580
Db      1807 ACAGAGCTCTACACTTCTACTGCTCAGCCATCTACGAGAGCAGTGCAGTACTCT 1866
Qy      581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnAlaTyrSerValArgArg 600
Db      1867 GCGTACAGCTGACCGAGCTGCGCACTTGCGAGAACCAATCTATCTTGCGCAGCGG 1926
Qy      601 GlnLysProSerAspArgAlaAspSerArgSerTrpHisGluGluSerProPheGlu 620

```

```

Db      1927 CAGAACCAAGTACAGAGCTGACTCGCGCGAGCTGGCATGAGAGACCCCTTTGAA 1986
Qy      621 LysGlnPheIleuArgArgSerCysGlnMetGluPheGlyGluSerIleuMetSerGluAsn 640
Db      1987 AAGCAGTTTAAAGCAGAGAGCTTCCAAATGGAATTTGGAGAGAGCATCATGTCAGAGAAC 2046
Qy      641 ArgSerArgGluIleuGluIleuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db      2047 AGGTACGGGAGAGAGCTGGGGAAGTGGCGATCATGCTTTCGGGACAGATGAA 2106
Qy      661 IleIleGluValSer 665
Db      2107 ATCATTGAGCTTCC 2121

RESULT 5
AEN59704
ID AEN59704 standard; cDNA; 3104 BP.
AC AEN59704;
XX
XX 28-JUN-2002 (first entry)
XX
XX Novel human coding sequence SEQ ID NO: 115.
XX
XX Human; antihaemic; vulnerary; antiinflammatory; immunomodulator;
XX antifertility; cerebroprotective; cyrostatic; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag; gene; ss.
XX Homo sapiens.
XX
XX WO200222660-A2.
XX
XX 21-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US026015.
XX
XX 11-SEP-2000; 2000US-00659671.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QH, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-292408/33.
XX P-PDB; ABB97291.
XX
XX An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis.
XX
XX Claim 1; SEQ ID NO 115; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX haemotopoesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a coding sequence of the
XX invention
XX
XX SQ Sequence 3104 BP; 845 A; 772 C; 718 G; 769 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,04e-234 Length: 3104
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: Gaps: 0

```

US-10-029-345A-109 (1-665) x ABNS9704 (1-3104)

QY 1 Metalahisglumecilleglythrghnillevalthrghuargleuvalaleuenglun 20
 Db 159 ATGGCCCATGAGATGATTTGGAACTCAATTTGTTACTGAGAGGTTGGCTGCTGGGAA 218
 QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 Db 219 AGTGAACGGAAGAAAGTGTCTAATTGATAGCCGCCCAATTTGTGGAAATACATACATCC 278
 QY 41 HisIleuGluIuAlaIleuIleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 Db 279 CACATTTTGGAAACCATTAATATCACTCTCCAGCTTATGAGCGAAGGTTGACATTC 338
 QY 61 AspIleValIleuIleThrGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
 Db 339 GACAAAGTGTATTATTCAGAGCTCATCCAGCATTCAGGAAACATMAGTTGACATTCAT 398
 QY 81 CysSerGluIleValValValValValValValValValValValValValValVal 100
 Db 399 TGCAGTCAGAGGTTGATGATTTACATCAAGCTCCCAAGATGTTGCTCTCTCTCTCA 458
 QY 101 AspCysPheLeuThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
 Db 459 GACTGTTTCTCAGCTGATCTTGGGTAACTGGAGAAAGCTTCAACTGTTCACTG 518
 QY 121 LeuValaglyGlyPheValGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
 Db 519 CTTCAGAGGTTGGTTTGGTGAATTTCTCTGATGTTTCTCTGCTCTGCTGAGAAATCC 578
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 Db 579 ACTCAGTCCCTACCTGATTCATTCAGACTTGTGCTTACCTGCTGCTGCTGCTGCTGCT 638
 QY 161 ArgIleuProAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
 Db 639 CGAATTTCTCCCAATCTTTATCTTGGCTGACAGAGATGCTCTCAACAGAGAGCTGATG 698
 QY 181 GlnIleAsnGlyIleGlyThrValLeuAsnAlaSerThrCysProLysProAspPhe 200
 Db 699 CAGCAGATGGATGGTATATGTTAATGCAAGCAATACCTGCTCCAAAGCTGACTT 758
 QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleu 220
 Db 759 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
 QY 221 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
 Db 819 CCTGGTGGACAAATCACTAGATTCATTGAGAAAGCAAAAGCTCCCATGATGATGTT 878
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerLysThrIleAlaIleAlaIleAlaIle 260
 Db 879 CTAGTGCACTGTTTATGCTGAGTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 938
 QY 261 LysArgMetAspMetSerLeuAspGluAlaValArgPheValLysGluLysAspProThr 280
 Db 939 AAGAGAGAGGACATGCTTTTATGATGAGCTTACAGATTTGTGAAGAAAAAGAACTTACT 998
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
 Db 999 ATATCTCCCACTTCATTTTCTGGGCCAACTCTGAGCATATGAGAAAGATTAAGAAC 1058
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisIleuGluLysProAsn 320
 Db 1059 CAGACTGGAGATCAGGGCCAAAGAGCAACTCAAGCTGCTGCACTGGAGAGCAAAAT 1118
 QY 321 GluProValProAlaValSerGluGlyGlyGlnLysSerGlnThrProLeuSerProPro 340
 Db 1119 GAACCTGTCTCTGCTGCTGAGGGTGGACAGAAAGGAGAACCCCTCAGTCAACCC 1178
 QY 341 CysAlaAspSerLysThrSerGluAlaAlaGlyIleArgProValHisProAlaSerVal 360

Db 1179 TGTGCCACTCTGCTACTCAGAGGCAAGCAAGAAAGCCCGTGCATCCCGCAGCTG 1238
 QY 361 ProSerValProSerValGlnProSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1298
 Db 1239 CCAGAGCTCCAGAGGTCAGAGCTGCTGTTTAAAGACAGCCGCTGTGATCAGAGGCTC 1298
 QY 381 SerGlyLeuHisIleuSerAlaAspArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 400
 Db 1299 AGTGGGCTGACCTGTCCGACAGAGCTGAGAGACAGCAATAGCTCAAGCTTCCCTTC 1358
 QY 401 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaIleSerLeuHisIle 420
 Db 1359 TCTCTGATATCAATCAATCAGTTTATATTCAGCCAGATGACAGCATCTTACATGCTTC 1418
 QY 421 SerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
 Db 1419 TCTCTATCAGAAAGTGTCTTGGAAATCTCAAACTTCCACTACTGTGATGGACCAAC 1478
 QY 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
 Db 1479 AAGCTATGCAAGTTCTCCCTGTTTCAAGAACTATCGAAGACATCCGAAACAGTCTCT 1538
 QY 461 AspLysGluGlnPheSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 Db 1539 GATTAAGAGAAAGCCAGCATCCCAAGAGCTGACAGACCGCAGGCTTTCAGACAGCCAG 1598
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 Db 1599 AGCAACGATTCATTCGCTGAGAACAGAGCAAGCTGACCCGCAAGAGTCTCTTTTAA 1658
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
 Db 1659 TCTCCACTGATCAGAAAGTGGAGGCTGGAGGCAATTAACACACAGCTTCTTTTCGG 1718
 QY 521 LeuSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheHis 540
 Db 1719 CTTCACACAGGCAAGCAGACCTCAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1778
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
 Db 1779 TCGGATATCTTGGCCCCCCCCAGACTCTTACCTCTCCCTGACCGAGCTGATTTTGGC 1838
 QY 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
 Db 1839 ACAAGATCCCTACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1898
 QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
 Db 1899 GCTACAGCTGCAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1958
 QY 601 GlnLysProSerAspAlaArgAlaAspSerArgArgSerTyrPheIleGluGluSerProPheGlu 620
 Db 1959 CAGAAAGCAAGTACAGAGACTGACTGCGGGAGCTGCAATGAGAGAGGCCCTTGAA 2018
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
 Db 2019 AAGCACTTTAAACGCAAGAGCTGCAAAATGGAATTTGAGAGACATCATGTGAGAAC 2078
 QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerArgLysMetGlu 660
 Db 2079 AGGTCAACGGGAAGACTGGGGAAGTGGCAGTCAAGTCACTTTTGGGCAAGATGAA 2138
 QY 661 IleIleGluValSer 665
 Db 2139 ATCATTTGAGGTCTCC 2153

RESULT 6
 ABK47596
 ID ABK47596 standard; cDNA; 3496 BP.
 XX ABK47596;
 AC
 XX
 DT 02-JUL-2002 (first entry)

XX cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.
DE
XX Human, dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KM mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KM cancer; graft-versus-host disease; allergy; metabolic disease;
KM abnormal cell growth; abnormal cell proliferation; contact inhibition;
KM cell cycle abnormality; anchorage independent cell growth; apoptosis;
KM intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; ss.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH CDS 562..2559
FT /tag= a
FT /product= "Human dual-specificity phosphatase-3 (DSP-16)
FT protein"
XX MO200226997-A2.
XX 04-APR-2002.
XX 25-SEP-2001; 2001MO-US030124.
XX 26-SEP-2000; 2000US-0235487P.
XX (CEPT-) CEPTIR INC.
XX Luche RM, Wei B;
XX P-PSDB; AAU9156.
XX WPI; 2002-315802/35.
XX DR P-PSDB; AAU9156.
XX New DSP-16 polypeptide, useful for identifying modulators of its
PT activity, which can be used in the treatment of disorders such as
PT Duchenne muscular dystrophy, or cancer.
XX
XX Claim 7; Fig 1; 87pp; English.
XX
XX The present invention relates to a new polypeptide, DSP-16, having a 665
CC amino acid sequence, given in the specification, or a variant having at
CC least 50 % identical residues, which retains the ability to
CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
CC invention can be used for identifying agents which modulate DSP-16
CC activity, for modulation of a proliferative response in a cell, survival
CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne
CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
CC cell proliferation, and cell cycle abnormalities. The present nucleic
CC acid sequence is that of the human dual-specificity phosphatase-3 (DSP-
CC 16) gene located on chromosome 12p. This sequence encodes the human DSP-
CC 16 protein of the invention
XX
SQ Sequence 3456 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,21e-234 Length: 3496
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: 6 Gaps: 0

US-10-029-345A-109 (1-665) X ABK47596 (1-3496)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 562 ATGGCCATGAGATGATTGGAACCAATTGTTACTAGAGAGTTGGTGGCTCTGCTGGAA 621
QY 21 SerGlyThrGluysValLeuLeuIleAspSerArgProPheValGluTyraAntHisSer 40

Db 622 AGTGAACGGAAGAAAAGTGTCTAATGATAGCCGCAATTGTGGAAATCAATATATCC 681
QY 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerIleuMetIleValArgLeuGlnGln 60
Db 682 CACATTTTGAAGCCATTAAATATCAATGCTCCAAAGCTTATGAGCAAGAGTTGCAACG 741
QY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleValHisValAspIleAsp 80
Db 742 GACAAAGTGTAAATTACAGAGCTCATTCACCATTCACCGAAACATTAAGTTGACATTGAT 801
QY 81 CysSerGlnIleValValValIleTyraSPGlnSerSerGlnAspValIleSerLeuSerSer 100
Db 802 TGCAGTGAAGAGTTGATGATTAGATTAGATCAAAAGCTCCCAAGATGTGCTCTCTTCA 861
QY 101 AspCysPheLeuThrValIleLeuGlyIleLeuGlyIleSerPheAsnSerValHisLeu 120
Db 862 GACTGTTTCTCACTGATCTTGGTAACTGGAGAAAGAGCTTCACTGTTCACTG 921
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuCysGlyIleValSer 140
Db 922 CTTCAGAGTGGTTTGTGAGTTCTCTGTTGTTCCCTGACCTCTGTAAGAAATCC 981
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 982 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTTGCCAACTGGGCCAAC 1041
QY 161 ArgIleLeuProAsnLeuTyraLeuGlyCysGlnArgAspValLeuAsnIleGlyLeuIle 180
Db 1042 CGAATTCCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCCCAAGAGAGCTATG 1101
QY 181 GlnGlnAsnGlyIleGlyTyraIleAsnAlaSerTyraThrCysProIleProAspPhe 200
Db 1102 CACGAGATGGATGTTATGTTATGTTAAATGCCAGCAATCTGTCAAAGCCGACTTT 1161
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlyIleLeu 220
Db 1162 ATCCCGAGTCTATTTCCGCGTGTGCTGTAATATACAGCTTTGTGAAATTTTG 1221
QY 221 ProTrpLeuAspIleSerValAspPheIleGluValAlaIleAsnGlyCysVal 240
Db 1222 CCGTGTGTGACAAATCAGTAGATTTTCAATTAAGCAAAAGCTTCAATGATGTGTT 1281
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyraIleMet 260
Db 1282 CTAGTGAAGTGTAGTGGAGATCTCCGCTCCGACCAATGCTATGCTTACATATG 1341
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyraArgPheValIleGlyIleArgProThr 280
Db 1342 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 1401
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyraGluTyraIleValAsn 300
Db 1402 ATATCTCCAAATCTTAATTTTCTGGGCAACTCTGAGCATATGAAAGAAAGATTAAGAC 1461
QY 301 GlnThrGlyAlaSerGlyProIleSerIleLeuIleuMetIleGlnIleValProAsn 320
Db 1462 CAGACTGAGCAATCAGGCGCAAAAGAGCAAACTAAGCTGCTGCACTGGAGAAAGCCAAAT 1521
QY 321 GluProValProAlaValSerGlyIleGlyIleSerGluThrProLeuSerProPro 340
Db 1522 GAACCTGTCCCTCTCTGTCTAGAGGTTGACAAAGAAAGCAAGAGCCCTCACTCCACC 1581
QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
Db 1582 TGTGCCACTCTGCTACTCAGAGGAGGAGCAAGAGCCCTGATCCCGCAGAGGTG 1641
QY 361 ProSerValProSerValGlnProSerIleuGluAspSerProIleuValGlnAlaLeu 380
Db 1642 CCCAGGCTGCCAGCGGAGCCGCTGCTTTAGAGACAGCCGCTGGTACAGGCGCTC 1701
QY 381 SerGlyLeuHisIleuSerAlaAspArgLeuGluAspSerAsnIleuTyraIleValSerPhe 400

XW	acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers 589..2586 /*tag= a /product= "Dual specificity phosphatase 21117"
PN	
PX	MO200173059.A2.
PD	
XX	04-OCT-2001.
PF	
XX	23-MAR-2001; 2001WO-US009477.
PR	
XX	24-MAR-2000; 2000US-0191858P.
PA	(MILL-) MILLENNIUM PHARM INC.
PI	Meyers RA;
XX	
DR	WPI; 2001-611635/70.
DR	P-PsDB; AAU05016.
PT	New human dual specificity polypeptides and nucleic acids for diagnosis of disease and treatment of e.g. liver disorders.
PS	Claim 1; Fig 1; 143pp; English.
XX	
CC	The invention relates to two novel human dual specificity phosphatases designated 21117 and 38692, the nucleic acids encoding them (including fragments, allelic variants, their complements or nucleic acids that hybridize to them) and antibodies raised against the proteins. The antibody is useful for detecting the presence of the polypeptide, and the nucleic acid fragments are useful for detecting the presence of the nucleic acid, for diagnosing disease. Agents (e.g. the antibody and antisense sequences) are useful for modulating the activity or expression of the polypeptide/nucleic acid, useful for the treatment of e.g. hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus, rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others listed in the specification) liver disorders, erythroid associated disorders (e.g. haemolytic anaemia) cellular proliferative or differentiative disorders, leukemias (e.g. acute myeloid leukaemia), carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and 38692 are also useful for modulating the proliferation, survival and migration or differentiation of a 38692 or 21117-expressing cell. The polypeptide and nucleic acids are useful for identifying modulating agents. The present sequence encodes the dual specificity phosphatase 21117
XX	
SQ	Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
Alignment Scores:	
Prod. No.:	1,24e-234
Score:	3406..00
Percent Similarity:	99.85%
Best Local Similarity:	99.70%
Query Match:	99.65%
Gaps:	0
Length:	3544
Matches:	663
Conservative:	1
Mismatches:	1
Indels:	0
Gaps:	0

Alignment Scores:
Prof vs

freq. NO.:	1,246-234	Length:	3544
Score:	346.00	Matches:	663
Percent Similarity:	99.85%	Conservative:	1
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	99.65%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-109 (1-665) X AAS14639 (1-3544)

1 METALAHISGLMETILEGLYThrglnIleValThrglnArgLeuValAlaLeuLeuGlu 20

-----CAGAACACAAATGTACTGAGAGGTTGGTGGCTCTGCTGAA 648

[illegible]

QY 41 HisIleuGluAlaIleasnIleasnCySSerLysIleuMetIvsArcaI aucI acI - 50

709 CACATTTTGGAGCCATTATATCACTGCTCCAAGCTTATGAAGCGAAGTTGCAACAG 768

QY 61 AsplvValLeuileThrgluleuileglnHlserAlalysHlslyValaAsp 80
 DB 769 GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGCAAAAGGTTGACATTGAT 828
 QY 81 CysSerGlnlyValValValValAspGlnSerSerGlnAspValAlaserLeuSer 100
 DB 829 TGCAGTCAGAAAGGTGTAGTTAGCATCAAGGCTCCCAAGATGTGGCTCTCTCTTCA 888
 QY 101 AspCyPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHlsleu 120
 DB 889 GACTGTTTCTCACTGACTTCTGGGTAACTGGAGAAAGACTTCACTGTTCCCTG 948
 QY 121 LeuAlglVglYpheaLaglpheAspSerArgCyPheProGlyLeuCyGglGlyLysSer 140
 DB 949 CTTCGAGGTGGGTTTGGTGGATCTCTGTTGTTTCCCTGGCTGTGAGAAATCC 1008
 QY 141 ThrLeuValProThrCyAlaserGlnProCyLeuProValAlaserleGlyProthr 160
 DB 1009 ACTCTAGTCCCTACCTGCAATTTCTGAGCCTTGCTTACCTTGTGCAATGGGGCCAAAC 1068
 QY 161 ArgIleLeuProAsnLeuTyLeuGlyCyeglnArgAspValLeuAsnlyGlyLeu 180
 DB 1069 CGAATTCCTCCCAATCTTATCTTGGGCTGCAGAGATGCTCTCAACAAGAGCTGATG 1128
 QY 181 GlnGlnAsnGlyLleGlyTyrrValLeuAsnAlaserTyrrThrCyProlyProAspPhe 200
 DB 1129 CAGCAGAAATGGAGTGGTATGTGTTAAATGCAAGCAATACCTGTCMAAGCTGACTTT 1188
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCyGglLysIleLeu 220
 DB 1189 ATCCCGAGTCTATTTCCGCGTGTGCCCTGTAATGACAGCTTTGTGAAATTTTG 1248
 QY 221 ProTrpLeuAspLysSerValAspPheleGlyLysAlalysAlaserAsnlyCyVal 240
 DB 1249 CCTGTGTGCAAAATCAGTAGATTTCTATGAGAAAGCAAAACCTCAATGATGATT 1308
 QY 241 LeuValHisCyAlaserLeuLaglyLleSerArgSerAlaThrleAlalaserleuMet 260
 DB 1309 CTAGTGCATGTTAGTGGGATCTCCGCTCCGCAACATCCCTACCTACATCATG 1368
 QY 261 LysArgMetAspMetSerLeuAspGlnAlaTyrrArgPheValLysGlnLysArgProthr 280
 DB 1369 AAGAGGATGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAACCTACT 1428
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuAsnAspTyrrLysLysIleLysAsn 300
 DB 1429 ATATCTCAAACTTCAATTTCTGGGCAACTCTGCACTATGAGAAAGATTAAAGAAC 1488
 QY 301 GlnThrGlyAlaserGlyProlySerLysLeuLeuLeuHlsleuGlyLysProAsn 320
 DB 1489 CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAAGCTGTGACCTGGAGAAACCAAAAT 1548
 QY 321 GluProValProAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerProPro 340
 DB 1549 GAACCTGTCCCTCTGTCTCAGAGGTTGACAAAGAAAGCAAGCGCCCTCAATCCACC 1608
 QY 341 CyAlaserSerAlaThrSerGlnAlalaglyGlnArgProValHisProAlaserVal 360
 DB 1609 TGTGCCGACTCTCTCACTCAGAGGAGCAGAGCAAAAGGCGCTGATCCCGCAGAGT 1668
 QY 361 ProSerValProSerValGlnProSerLeuLeuLysAspSerProLeuValGlnAlaLeu 380
 DB 1669 CCGAGGCTGCCAGCGTGCAGCGCTGTGTAGAGACAGCCCGCTGTGACAGGGCTTC 1728
 QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlyLysSerAsnLysLeuLysArgSerPhe 400
 DB 1729 AGTGGGCTGACCTGTCCGAGACAGGCTCGAAGACAGCAATAGCTCAAGCTTCTTC 1788
 QY 401 SerLeuAspLleLysSerValSerTyrrSerAlaserMetAlalaserLeuHlslyPhe 420
 DB 1789 TCTCTGATATCAAAATCAGTTTCAATTTACGACAGATGGCAGCATCTTACATGCTTC 1848

QY 421 SerSerSerGlnAspAlaLeuGlnTyrrTyrrProSerThrThrLeuAspGlyThrAsn 440
 DB 1849 TCTCTATCAGAGATGCTTTGGAAATCTACAACTTCCATCTCTGATGGAGCAAC 1908
 QY 441 LysLeuCyeglnPheSerProValGlnLeuSerGlnGlnThrProGlnThrSerPro 460
 DB 1909 AAGCTATGCCAGTTCTCCCTGTTAGAAATCTACGAGAGACTCCGCAACAGCTCT 1968
 QY 461 AsplvGlyGlnAlaserleProlyLysLeuGlnThrAlaProSerAspSerGln 480
 DB 1969 GATTAAGAGAGAACCCAGCATCTCCCAAGAGCTGCAACCGCGGCTTTAGACAGCAG 2028
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 DB 2029 AGCAGCGATTGATTCGGTCAGAACCAAGAGAGTGGACCCGCCAGAGCTCCTTTTA 2088
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTyrrHisThrSerPheLeuPheGly 520
 DB 2089 TCTCAGCTGATCGAAGTGGAGCGTGAAGACAAATTAACACAGACTTCTTTTCGGC 2148
 QY 521 LeuSerThrSerGlnGlnHlsleuThrLysSerAlaglyLeuGlyLeuLysGlyTrpHis 540
 DB 2149 CTTTCCACAGCCAGACAGCACTCAAGAACTGTGCTGGGCTGTAAAGGCTGGCAC 2208
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTyrrPheAla 560
 DB 2209 TCGGATATCTTGGCCCCCAGACCTTACCTTCCCTTGACAGAGCTGTATTTTGGC 2268
 QY 561 ThrGlnSerSerHisPheTyrrSerAlaserAlaIleTyrrGlySerAlaserTyrrSer 580
 DB 2269 ACAGATCTCTCACTTACTGCTGCTCAGGCATCTAAGAGAGCGAGTGCATTTACTCT 2328
 QY 581 AlaTyrrSerCySerGlnLeuProthrCyeglyAspGlnValTyrrSerValaArgArg 600
 DB 2329 GCTTACAGCTGCAGCGACCTGCCACTTGGGAGCAAGTCTATTTCTGTGCCAGCGG 2388
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGlnGlnSerProPheGln 620
 DB 2389 CAGAACCAAGTACAGAGCTACTGCGCGGAGCTGGCATTAAGAGACCCCTTTGAA 2448
 QY 621 LysGlnPheLysArgArgSerCyGlnMetGluPheGlyGlnSerIleMetSerGluAsn 640
 DB 2449 AAGCAGTTTAAACGAGAGCTGCCAAATGAAATTTGAGAGAGCATCATGTCAAGAAC 2508
 QY 641 ArgSerArgGlnGlyLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 DB 2509 AGGTCAAGGAGAGCTGGGAGAAAGTGGGACAGTCACTTTTCTGGGACACATGAA 2568
 QY 661 IleIleGluValSer 665
 DB 2569 ATCATTTAGAGTCTCC 2583

RESULT 8
 ID ABK49402
 AC ABK49402;
 DT 02-JUL-2002 (first entry)
 DE cDNA encoding human dual specificity phosphatase 21117 protein.
 XX Human; dual specificity phosphatase 21117; erythroid-related disorder;
 KW haematopoietic-related disorder; leukemia; autoimmune disease; anaemia;
 KW erythrocytosis; liver-related disorder; cancer; gene; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 589..2586
 FT /tag= a
 FT /product= "Human dual specificity phosphatase 21117"
 FT /note= "Specifically claimed in claim 2"

XX US2002034807-A1.
 PN
 XX
 PD 21-MAR-2002.
 XX
 XX 23-MAR-2001; 2001US-00816494.
 PF
 XX
 PR 24-MAR-2000; 2000US-0191858P.
 XX
 PA (MEYE/) MEYERS R A.
 XX
 PI Meyers RA;
 XX
 DR WPI: 2002-351088/38.
 DR P-PSDB; AAU79929.
 XX
 PT New nucleic acids, designated 38692 and 21117, encoding dual specificity
 PT phosphatases for treating cell proliferation and differentiation
 PT disorders including hematopoietic and erythroid-related disorders and
 PT cancers.
 XX
 PS Claim 2; Fig 1; 76pp; English.
 XX
 CC The present invention relates to new nucleic acids designated 38692 and
 CC 21117 encoding dual specificity phosphatase family members. The nucleic
 CC acid, polypeptide encoded by it, and antibody specific for the
 CC polypeptide may be used to diagnose and treat haematopoietic-related
 CC disorders such as leukemias and autoimmune diseases, erythroid-related
 CC disorders such as anemias and erythrocytosis, liver-related disorders,
 CC and cancers, particularly of the breast, colon, adipose, prostate and
 CC lung. The present nucleic acid sequence encodes the human dual
 CC specificity phosphatase 21117 protein of the invention, as described
 CC above
 XX
 SQ Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. NC.: 1.24e-234 Length: 3544
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: Gaps: 0
 US-10-029-345a-109 (1-665) x ABK49402 (1-3544)
 QY 1 MetAlHisGluMetIleGlyThrGluIleValThrGluArgLeuValAlaIleuLeuGlu 20
 DB 589 ATGGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGTTGGTGGCTTGGTGGAA 648
 QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 DB 649 AGTGGACCGGAAAAGCTGCTGTAATTGATAGCCGGCCATTGGGAATCAATACATCTCC 708
 QY 41 HisIleuGluAlaIleAsnIleAsnCySerSerLeuMetLeuAspArgLeuGln 60
 DB 709 CACATTTTGGAAAGCATTAATATCACTGCCCAAGTTTGAAGGAGAGGTGGCAACG 768
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisValAspIleAsp 80
 DB 769 GACAAAGTTTAATTAAGAGCTCATCCAGCATTCACGAAACATAGGTGACATTAGT 828
 QY 81 CysSerGlnLysValValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSer 100
 DB 829 TGCAGTCAGAAAGTTGATGTTACGATCAAAAGCTCCCAAGATTTGCTCTCTCTTCA 888
 QY 101 AspCysPheLeuThrValIleuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
 DB 889 GAGCTTTTCTACTGTACTTCTGGGTAAACTGGAGAACGCTTCAACTCTGTTCACTG 948
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
 DB 949 CTTCAGAGTGGTCTGCTGATGTTCTCTGTTTCCCTGGCTCTGTGAAGGAAATCC 1008

QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 1009 ACTTAGTCCCTACCTGATTTTCTCAGCTTGTCTTACTCTGTGGCAACATTTGGCCACCC 1068
 QY 161 ArgIleuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuIle 180
 DB 1069 GAAATCTTCCCAATCTTATCTTGGCTGCCAGGAGATGCTCTCAACAGAGCTGATG 1128
 QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
 DB 1129 CAGCAGATGGGATTTGGTATGTTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 1188
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
 DB 1189 ATCCCGAGTCTCAATTTCTCGCGTGGCTGTGAATGACAGCTTTTGTGAGAAAATTTTG 1248
 QY 221 ProTyrLeuAspLysSerValAspPheIleGluValAlaLysValAsnGlyCysVal 240
 DB 1249 CCGTGTGGCAAAATAGTATGATTTCAATTGAGAAAAGCAAAAGCTCCAAATGATGTGT 1308
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 DB 1309 CTAGTGACCTGTTTACTGTGGATCTCCGCTCCGCCACATGCTATGCTTACATCATG 1368
 QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
 DB 1369 AAGAGATGAGACATGCTTTATATAGACTTACAGATTTGGAAGAAAAGAAAGCTTACT 1428
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
 DB 1429 ATATCTCCAACTTCAATTTCTGGCCCACTCTGACATGATGAGAAAGATTAGAAC 1488
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuIleGluLysProAsn 320
 DB 1489 CAGACTGGAGCATCAGGGCCAAAGAGCAACTCAAGCTGTGACCTGGAGAAAGCCAAAT 1548
 QY 321 GluProValProAlaValAsnSerGlyGlyGlnLysSerGlnThrProLeuSerProPro 340
 DB 1549 GAACTGTCCCTGCTGCTGTCTCAGAGGTGACAGAAAAGCAAGCCCTCGTCAAGTCAACC 1608
 QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyAlaArgProValHisProAlaSerVal 360
 DB 1609 TGTGCGCACTGTGCTACTCAGAGGACAGAGCAAAAGCCGTGATCTCCGACGCTG 1668
 QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 DB 1669 CCCAGCTGCCCAAGCGTGCAGCGCTGTGTTAGAGGACAGCCCGCTGTACAGGCGCTC 1728
 QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
 DB 1729 AGTGGGCTGCACCTGTCCGAGACAGGCTGGAAAGACAGCATTAAGCTTAAGCTTCTC 1788
 QY 401 SerLeuAspIleLysSerValSerTyrSerAlaSerMetValAlaSerLeuHisGlyPhe 420
 DB 1789 TCTCTGATATCAAAATCAATTTTCAATTCAGCCAGCATGAGAGATCTTACATGGCTTC 1848
 QY 421 SerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
 DB 1849 TCTCTATCAGAAAGATGCTTTGGATATCAAAACCTTCACTATCTGAGATGGAGCAAC 1908
 QY 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGlnGlnThrProGluThrSerPro 460
 DB 1909 AAGCTATGCAATTTCTCCCTGTTCAGGAATTCATCGAGAGAGACTCCGCAACCAATGCT 1968
 QY 461 AspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerArgSerGln 480
 DB 1969 GATTAAGGAGAAAGCAGCATCCCAAGACCTGAGACCCGAGGCTTTCAGACACCCAG 2028
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 DB 2029 AGCAAGCATTTGATTCGGTGAAGAACAGAGAGATGAGACCGCCAGAGATCCCTTTTA 2088

QY 501 SerProteinLeuArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
DB 2089 TCTCCACTGATCGAAGTGGAGCGTGGAGCAATTACACACACACTTCTTTGGC 2148
QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuGlyTyrHis 540
DB 2149 CTTTCACACGACGACGACACTCAGAGAGTGTGCTGGGCTTAAAGGCTGGCAC 2208
QY 541 SerAspLysLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
DB 2209 TCGGATATCTTGACCCCGACACTTACCCCTTCCCTGACACGACACTGTATTTGGC 2268
QY 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
DB 2269 ACAGAGTCTTCACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2328
QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
DB 2329 GCTTACAGCTGACGACGACGCTGCCACTTGGGAGACCAAGTCTATTTCTGCGCAGGCGG 2388
QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrPheGluGluSerProPheGlu 620
DB 2389 CAAAGGCAAGTACAGAGCTGACTCGCGCGGAGCTGATGAGAGACCCCTTTGAA 2448
QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
DB 2449 AACGAGTTTAAACGACAGAGCTGCCAATGGAATTTGAGAGACATCATGTCAGAGAC 2508
QY 641 ArgSerArgGlnGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
DB 2509 AGGTCAAGGAGAGAGCTGGGGAAGTGGGAGTCAAGTCAAGTCTTTGGGCGACAGATGAA 2568
QY 661 IleIleGluValSer 665
DB 2569 ATCATTTAGAGTCTCC 2583
RESULT 9
ID ABK14474 standard; cDNA; 3766 BP.
AC ABK14474;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human protein phosphatase 7 (PP7) cDNA sequence.
XX
KW Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;
KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
KW melanoma; myeloma sarcoma; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 536..2535
FT /tag= a
FT /product= "Protein_phosphatase_7_(PP7)"
XX
PN MO200210363-A2.
XX
PD 07-FEB-2002.
XX
PF 26-JUL-2001; 2001MO-US023716.
XX
PR 28-JUL-2000; 2000US-0221679P.
PR 03-AUG-2000; 2000US-0223272P.
PR 10-AUG-2000; 2000US-0224309P.
PR 18-AUG-2000; 2000US-0226728P.
PR 30-AUG-2000; 2000US-0229254P.
PR 08-SEP-2000; 2000US-0231366P.
XX

PA (INCY-) INCYTE GENOMICS INC.
XX
PI Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;
PI Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalia AJL, Lu DAM,
PI Tribouley CW, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
PI Walla NK, Kearney L;
XX
DR WPI; 2002-188735/24.
DR P-PSDB; AAU75789.
XX
PT New protein phosphatases, useful for diagnosing, treating or preventing
PT immune system disorders (e.g. Crohn's disease), neurological disorders
PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
PT cancers).
XX
PS Claim 5, Page 114-115, 117pp; English.
XX
CC The present invention relates to a new polypeptide, a naturally occurring
CC amino acid sequence at least 95 % identical to it, a biologically active
CC fragment of it or an immunogenic fragment of it. The polypeptides,
CC polynucleotides, agonists and antagonists are useful for diagnosing,
CC treating or preventing disorders associated with aberrant expression of
CC protein phosphatases (PP), particularly immune system disorders e.g.
CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
CC asthma or Crohn's disease, neurological disorder e.g. epilepsy,
CC Huntington's disease, dementia or Parkinson's disease, developmental
CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
CC or sarcoma. The present nucleic acid sequence encodes human protein
CC phosphatase 7 (PP7) which is one of several human protein phosphatases
CC (AAU75789-AAU75792) of the invention
XX
SQ Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.34e-234 Length: 3766
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: Gaps: 0
US-10-029-345A-109 (1-665) x ABK14474 (1-3766)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 538 ATGGCCCATGAGATGATTTGGAACTCAAAATTTGATCTGAGAGTGTGGCTGCTGGAA 597
QY 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
DB 598 AGTGGAAACGAAAGATGCTGCTAATTTGATGAGCGGCATTGTGGAAATACATACATCC 657
QY 41 HisIleLeuGlnValIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 60
DB 658 CACATTTTGGAGCCATTATATNCACTGCTCAACTTATGAGGAGGAGGTTGCAACAG 717
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAlaPheAsp 80
DB 718 GACAAAGTATTATTACAGAGCTCATTACAGCATTCAGCGAAATATGATGATTCATTGAT 777
QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 778 TGCAGTCAGAGAGTGTAGTTTACATCAAGATCCCAAGATGTGTCTCTCTCTTCA 837
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
DB 838 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAAGAGAGGTTCAACTCTGTCACTG 897
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGluGlyLysSer 140
DB 898 CTTCGAGGTGGGTGCTGATGTTCTCTGTTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 957
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160

Db 958 ACTCTACTCCCTTACCTGCATTTCTCAGCCCTTGCTTACCTGTGGCAACATTTGGCCGACCC 1017

QY 161 ArgIIeLeuProAnleuTYrIleuGIYcysGlnArgApValIleuAnlySGIleuIle 180

Db 1018 CGAATTCCTCCCAATCTTATCTTGGCGTCCAGGCAATGCTCTCCACAAAGAGAGCTCATG 1077

QY 181 GlnGlnAnsgIYIleGIYrValIleuAnlnAseTYrThrCysProIysProAspPhe 200

Db 1078 CAGCAGAAATGGAGATTGGTATGTGTTAAATCCAGCAATACCTGTCCAAAGCTCACTT 1137

QY 201 IleProGluSerHisPheIleuArgValProValAsnAspSerPheYsgIuIysIleu 220

Db 1138 ATCCCGAATCCATTTCTCTCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1197

QY 221 ProTProIleuAspIysSerValAspPheIIeGIuYalAlaYalAseAsnGIYcysVal 240

Db 1198 CCGTGGTTGGCAAAATCGATAGATATTTATGTGAAAGAAAGAAAGCCCTCCAAATGATGCTT 1257

QY 241 IeuValHisCysIleuIaGIYIleSerArgSerAlaThrIleAlaIIeAlaTYrIleMet 260

Db 1258 CTATGGCACTTTTATGGGATCTCCCGTCCGCAACATGCTATGCTCTACATCATG 1317

QY 261 IysArgMetAspMetSerIleuAspGluAlaIYrArgPheValYsgIuIysArgProIhr 280

Db 1318 AAGAGATGAGCATGATCTTTAGATGAAGCTTCACAATTTGTGAAGAAAGAAATTAAGAAC 1377

QY 281 IleSerProAnPheAnPheAnPheIleuGIYIleuIleuAspIYrGIYIysIleYsaAn 300

Db 1378 ATATCTCCAAATCTCAATTTTCTGGGCCAACTCTCGACCTAGCATGTGAAAGAAATTAAGAAC 1437

QY 301 GlnThrGlyAlaSerGIYProIysSerIysIleuIysIleuIleuIleuIleuIleuIleu 1437

Db 1438 CAGACTGGACATCAAGGCGCAAGGCAAACTCAAGCTGTGCACCTGGAGAAAGCCAAAT 1497

QY 321 GluProValProAlaValSerGIuIYGIYGIuIysSerGIuIhrProIleuSerProPro 340

Db 1498 GAACCTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAGCCAGAGCGCTTCAGTCAACC 1557

QY 341 CysAlaAspSerAlaThrSerGIuAlaIaGIYGIuArgProValHisProAlaSerVal 360

Db 1558 TGTGGCGACTGTCTCACTCAAGAGCGAGCGAGCAAAAGCCGTGATCCGCGCAAGCTG 1617

QY 361 ProSerValProSerValGlnProSerIleuGIuAspSerProIleuValGlnAlaIleu 380

Db 1618 CCGAGGTGCCAGCGGTGCAGCTCGCTTGTAGAGACAGCCCGCTGTCAGAGCCCTC 1677

QY 381 SerGIYleuHisIleuSerAlaAspArgIleuGIuAspSerAsnIYIleuIYArgSerPhe 400

Db 1678 AGTGGGCTGCACCTGTCCGAGACAGCTGAGAGACAGCAATTAAGCTCAAGCTTCTTC 1737

QY 401 SerIleuAspIleYserValSerTYrSerAlaSerMetAlaIaIaSerIleuHisGIYpHe 420

Db 1738 TCTCTGGATATCAATCAATCAATTCATATTCAGCGACAGATGGACGATCTTAACTGCTTC 1797

QY 421 SerSerSerGIuAspAlaIleuGIuTYrTYrIYsProSerThrThrIleuAspGIYThrAn 440

Db 1798 TCTCTCATCAAGATGTGTTGGAAATCAAAACCTTCACTACTCGATGGAGAACCAAC 1857

QY 441 LysIleuCysGlnPheSerProValGIuIleuIleuSerGIuIhrIhrProGIuIhrSerPro 460

Db 1858 AAGCTATAGCCAGTTCCTCCCTGTTCAGAACTATCGAGAGACTCCCAAAACAGATCCT 1917

QY 461 AspIYsgIuGlnIleuIleuSerIleProIYsIleuGIuIhrIhrIhrIhrIhrSerPro 480

Db 1918 GATTAAGAGAGAACCCAGCATCCCCAAAGAGCTGCAGACCGCCAGGCTTCAACAGCCAG 1977

QY 481 SerIYsArgIleuHisSerValArgThrSerSerSerGIYThrAlaGlnArgSerIleu 500

Db 1978 AGCAAGCCGATTCATTCGTCAGAAACAGACAGAGTGGCACCGCCCAAGAGGCTCCCTTTA 2037

QY 501 SerProIleuHisArgSerGIYSerValGIuAspAsnTYrHisThrSerPheIleuPheGIY 520

QY	Db	2038	TCCTCACTGCATCGAATGGTGGAGCGCTGGAGAGCAATTACACACACAGCTTCTTTGGCC	2097
QY	Db	521	LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuGlyLysPheHis	540
QY	Db	2098	CTTTCACACACCGACGACGACCTCAGCAGAGCTCTGCTGCTGGCTTTAAGGGCTGGCAC	2157
QY	Db	541	SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpYrPheAla	560
QY	Db	2158	TCGGATATCTTGGCCCCCGACAGCCTCTACCCCTTCCCTCCGACAGCAGCTGTATTTTGGC	2217
QY	Db	561	ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer	580
QY	Db	2218	ACAAAGTCCCAACACTCTACTCTGCTGCTCAGCATCTTACGAGAGCAGTCCAGTTACTCT	2277
QY	Db	581	AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg	600
QY	Db	2278	GCTTACAGCTGCGACGACGCTGCGCCACCTTCGCGAGACCAAGTCTATCTGTGGCAGGCGG	2337
QY	Db	601	GlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSerProPheGlu	620
QY	Db	2338	CAGAACCCAACTGACAGACTGACTGCTGGCGGCGAGCTGGCATGAAGAGGCCCTTTGAA	2397
QY	Db	621	LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn	640
QY	Db	2398	AGCGACTTTAAACGACGAAGCTGCGCAATGGAAATTGGAGAGAGCATCATGTGACAGAGAC	2457
QY	Db	641	ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu	660
QY	Db	2458	AGGTACCGGGAAGAGCTGGGGAAGTGGGCGAGCTGACTTACCTTTTGGGCGAGCATGGAA	2517
QY	Db	661	IleIleGluValLysSer	665
QY	Db	2518	ATCATTGAGCTCTCC	2532
RESULT 10				
ID	ABN83966	standard; DNA; 4790 bp.		
AC	ABN83966;			
XX				
DT	06-SEP-2002	(first entry)		
XX				
DE	Human gene sequence #13.			
KW	Human; brain; tonsil; hippocampus; foetal brain; diagnosis; gene; ds.			
XX				
OS	Homo sapiens.			
XX				
PH	Key	Location/Qualifiers		
FT	CDS	184..2181		
XX		/tag= a		
PN	WO200252005-A1.			
XX				
PD	04-JUL-2002.			
XX				
PF	20-DEC-2001; 2001MO-JP011217.			
XX				
PR	22-DEC-2000; 2000JP-00389742.			
XX				
PA	(KAZU-) KAZUSA DNA RES INST FOUND.			
PA	(CELE-) CELESTAR LEXICO-SCI LTD.			
XX				
PI	Ohara O, Nagase T, Nakajima D;			
XX				
DR	WPI; 2002-500762/53.			
XX	P-P5DB; ABB97946.			
XX				
PT	Genes and their expression products cloned from human cDNA libraries for			
XX	treatment and diagnosis of diseases associated with their expression.			
XX	Claim 1(a); Page 111-117; 238bp; Japanese.			

Genes and their expression products cloned from human cDNA libraries for treatment and diagnosis of diseases associated with their expression.
Claim 1(a), Page 111-117, 238pp; Japanese.

CC The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies,
CC oligonucleotide probes and DNA chips for diagnosis and identification of
CC drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABN83954-ABN83984 represent
CC human gene sequences of the invention

XX Sequence 4730 BP; 1293 A; 1064 C; 1152 G; 1281 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.85e-234	Length:	4730
Score:	3406.00	Matches:	663
Percent Similarity:	99.85%	Conservative:	1
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	99.65%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345a-109 (1-665) x ABN83966 (1-4730)

Qy 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
Db 184 ATGGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGTTGGCTGCTGCTGAA 243
Qy 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyraenthSer 40
Db 244 AGTGAAGCGAAAGAGCTGCTAATGATAGCCGCGCATTTGGAAATCAATACATCC 303
Qy 41 HisIleLeuGlnIleAlaIleAsnIleAsnCysSerIleuMetIleuAspArgLeuGln 60
Db 304 CACATTTTGAAGCATTAATATCACTGCTCCAAAGCTTAAGAACGAAAGTTGCAAG 363
Qy 61 AspLeuValLeuIleThrGluLeuIleGlnHisSerAlaIleValIleAsp 80
Db 364 GACAAAGTATTAATACAGAGCTATCCACATTCACGAAACAAAGTGGACATTGAT 423
Qy 81 CysSerGlnIleValValIleValIleAspGlnSerSerGlnIleAspValIleSer 100
Db 424 TGCAGTCAGAAAGTGTGATGTTACGATCAAGCTCCAAAGATGTGCTCTCTCTTCA 483
Qy 101 AspCysPheLeuThrValIleuLeuGlyValLeuGlnIleuSerPheAsnSerValIle 120
Db 484 GACTGTTTCTCACTGACTTCTGGGTAACTGGAAGAGCTTCACTGTTCACTG 543
Qy 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGlyIleuSer 140
Db 544 CTTCGAGGTGGGTTTCTGAGGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 603
Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 604 ACTGATGCTCTACCTGATTTCTCAGCTTGTGCTTACCTGTGGCAACATTTGGCCAA 663
Qy 161 ArgIleuProAsnLeuIleuGlyCysGlnArgAspValIleuAsnIleuGlyLeuIle 180
Db 664 CGAATCTTCCCAATCTTATCTTGTGCTGACGAGATGCTCTCAACAAAGAGCTGATG 723
Qy 181 GlnGluAsnGlyIleGlyIleValIleuAsnAlaSerTyThrCysProIleuProAspPhe 200
Db 724 CAGCAAGATGGATGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCTTGACTTT 783
Qy 201 IleProGluSerHisPheLeuArgValIleValIleAsnAspSerPheCysGlyIleu 220
Db 784 ATCCCGAGTCTCATTTCTGCGTGTCTGTGAATACAGACTTTTGTGAAGAAATTTTG 843
Qy 221 ProTrpLeuAspIleuSerValAspPheIleGlyValIleValIleAsnIleuGlyCysVal 240
Db 844 CCGGTGTGGAACAATAGTATGATTTCAATGAGAAACAAAGCTTCAATGATGTTG 903
Qy 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260

Db 904 CTAGTCACTGTTTACGTGGATCTCCCGCTCCGCCACCATCGCTATCGCTATCATCATG 963
Qy 261 LysArgMetAspMetSerLeuAspGluAlaTyraPheAlaIleGlyIleuAspProThr 280
Db 964 AAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAAGACCTACT 1023
Qy 281 IleSerProAsnPheAsnPheLeuGlyIleuLeuAspTyrlleuValIleValAsn 300
Db 1024 ATATCTCCAACTTCAATTTTCTGGGCCCACTCTGACATATGAGAAAGATTAAGAAC 1083
Qy 301 GlnThrValAlaSerGlyProIleuSerIleuValLeuLeuIleuGlyIleuProAsn 320
Db 1084 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGACCTGAGAGAACCAAT 1143
Qy 321 GluProValProAlaValSerGlyIleuGlyIleuValSerGlyIleuProLeuSerProPro 340
Db 1144 GAACCTGCTCTGCTCTCTCAAGAGGTGACAAAGAGAGAGCGCTTCACTTCAACC 1203
Qy 341 CysAlaAspSerAlaThrSerGluAlaIleGlyIleuArgProValHisProAlaSerVal 360
Db 1204 TGTGCGGACTCTCTACCTCAAGAGCAGAGCAAGAGCCCTGTGATCCCGCAGCGTG 1263
Qy 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
Db 1264 CCGAGGTGCGCCAGCGGAGCGCGTGTGATGAGAGAGCGCGCTGTGATCAGGCGCTC 1323
Qy 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleuValArgSerPhe 400
Db 1324 AGTGGCTGCACTGCTCCGAGCAGAGCTGGAAGAGCAATAGCTTCAAGCTTCTTC 1383
Qy 401 SerLeuAspIleuSerValSerTyraSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
Db 1384 TCTTGATATCAAAATCATGTTATATTCAGCCAGATGACACATCTTAATGCTTC 1443
Qy 421 SerSerSerGlnAspAlaLeuGlyTyrlleuProSerThrThrLeuAspGlyThrAsn 440
Db 1444 TCTCATCAGAAAGATGCTTGGAAATCACTCAAACTTCCATCTGATGGAGCAAC 1503
Qy 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluIleuThrProGluThrSerPro 460
Db 1504 AAGCTATGCAAGTCTCCCTGTTAGAGAACTTTCAGAGCACTCCCGAAACAGACT 1563
Qy 461 AspLeuGluGluAlaSerIleProIleuValLeuGlnIleuThrAlaArgProSerAspSerGln 480
Db 1564 GATAGAGAGAGACCAAGCATCCCAAGAGCTGCAAGCCGACGCTTACAGACAGCAG 1623
Qy 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
Db 1624 AGCAAGCATGATTCGGTCAAGAACCAAGAGAGAGAGCGCCAGAGGTCCCTTTTA 1683
Qy 501 SerProLeuHisArgSerGlySerValGluAspAsnTyrlleThrSerPheLeuPheGly 520
Db 1684 TCTCCACTGATGAGAGTGGAGCGTGGAGAACTTACCAACAGACTTCTTTCGCG 1743
Qy 521 LeuSerThrSerGlnGlnHisLeuThrIleuValIleuGlyIleuValIleuValIleu 540
Db 1744 CTTTCAACAGCAGAGCACTTCAAGAGTCTGAGGCTGAGGCTTAAAGGCTGGCAC 1803
Qy 541 SerAspIleuAlaProGlnThrSerThrProSerLeuThrSerSerTyrllePheAla 560
Db 1804 TCGGATATCTGGCCCCCAGACCTTACCCCTTCTGACAGAGAGCTGTATTTTGGC 1863
Qy 561 ThrGluSerSerHisPheTyraSerAlaIleTyrlleGlyIleuSerAlaSerTyrlle 580
Db 1864 ACAGAGTCTCTCACTTCACTTCTGCTGAGCATCAGAGGAGAGGCTGCACTTCTCT 1923
Qy 581 AlaTyraSerCysSerGlnIleuProThrCysGlyAspGlnValTyraSerValArgAspArg 600
Db 1924 GCTTACAGCTGAGCGAGCTGCCACTTGGAGAGCAAGCTTATTTCTGCGCAGGCGG 1983
Qy 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrlleGluIleuSerProPheGlu 620
Db 1984 CAGAAAGCAAGTGAAGAGCTGATCTGCGCGGAGCTGGCATGAAAGAGCCCTTTGAA 2043

QY 621 LysGlnPheLysArgSerCysGlnMetGluPheGlyCysSerIleMetSerGluasn 640
 DB 2044 AAGCAGTTTAAACCCAGAGCTGCCAATGAAATTTGGAGAGCATCATGTCCAGGAAC 2103
 QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 DB 2104 AGGTCACGGGAGAGGCTGGGGAAAGTGGCAGTCACGTCTTTCCTTTCCGAGCAGATGAA 2163
 QY 661 IleIleGluValSer 665
 DB 2164 ATCATTGAGGTCTCC 2178
 RESULT 11
 ABV20833
 ID ABV20833 standard; cDNA; 5145 BP.
 XX
 AC ABV20833;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 20824.
 XX
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183119P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MIL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 3419; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (g)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: 2.03e-234 Length: 5145
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1

Query Match: 99.65% Indels: 0
 DB: 5 Gaps: 0
 US-10-029-345a-109 (1-665) x ABV20833 (1-5145)
 QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluValAlaLeuGlu 20
 DB 589 ATGCCCATGAGATGATGAACTCAATGTTTACTAGAGAGTTGGCTCTGCTGGAA 648
 QY 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluThrAsnThrSer 40
 DB 649 AGTGGAAAGGAAAAAGTGGCTGTAATTGATAGCCGGCATTTTGGAAATACATACATCC 708
 QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGln 60
 DB 709 CACATTTTGGAAAGCATTAATATCACTGCTCCAAAGCTTATGAAAGGAAAGTTGCAACG 768
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysValAspIleAsp 80
 DB 769 GACAAAGTTTAATTACAGAGCTCATCCAGCATTCACGAAACATPAGGTGACATTGAT 828
 QY 81 CysSerGlnLysValValValIleThrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 829 TGCAGTCAGAAAGGTGTAGTTTACGATCAAAAGCTCCAAAGATGTTGCTCTCTCTTCA 888
 QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlySerPheAsnSerValHisLeu 120
 DB 889 GACTGTTTCTCACTGTACTCTGGGTAACTCGAAGAAAGCTTCACTGTTCACTG 948
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
 DB 949 CTTCAGAGTGGGTTTGTGAGTTCCTGCTGTGTTTCCCTGGCCTCTGAAAGAAATTC 1008
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 1009 ACTCTAGTCCCTCACTGCACTTCAGCTTCTGCTTCTGTTGCAACATTTGGCCCAACC 1068
 QY 161 ArgIleLeuProAsnLeuThrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuIle 180
 DB 1069 CGAATTTCTCCAAATCTTATCTTGCTGCTGCAGAGATCTCTCAAAAGAGCTGATG 1128
 QY 181 GlnGlnAsnGlyIleGlyThrValLeuAsnAlaSerThrCysProLysProAspPhe 200
 DB 1129 CAGCAGATGGAGATGGTTATGTTTAATGCCAAGCAATCTGTCAAAAGCTCACTT 1188
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
 DB 1189 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 1248
 QY 221 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysValSerAsnGlyCysVal 240
 DB 1249 CCGGTGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCAAAGAGATGTTT 1308
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 DB 1309 CTNAGTCAGCTTTTACTGAGATCTCCGCTCCGACCAATCGCATATGCTTACATCATG 1368
 QY 261 LysArgMetAspPheSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
 DB 1369 AAGAGGATGAGATGCTTTAGATCAACTTACAGATTTTGGAAAGAAAGAAAGCTTACT 1428
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
 DB 1429 ATATCTCCAAACTTCAATTTCTTGCGCCAACTCTCGATCATATGAGAGAAATTGAAGAC 1488
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGluLysProAsn 320
 DB 1489 CAGACTGAGATCATCGGCAAGCAAACTCAAGCTCTGCACTCGGAGAAAGCCAAAT 1548
 QY 321 GluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeuSerProPro 340
 DB 1549 GAACCTGTCCTGCTGCTCAGAGGTGACAGAAAGGAGAGCCCTCAGTCCACCC 1608

```

QY 341 CyeAlaAapSerAlaThrSerGluAlaGlyGlnArgProValHisProAlaSerVal 360
DB 1609 TGCGCCAGCTCTGCTACTCAGAGCGACGACAAAGGCCGCTGATCCCGCCAGCGTG 1668
QY 361 ProSerValProSerValGlnProSerLeuGlnAapSerProLeuValGlnAlaLeu 380
DB 1669 CCCAGCGTGCCAGCGGCGAGCGCTGCTGTAGAGACAGCCGCGTGACAGCGGCTC 1728
QY 381 SerGlyLeuHisLeuSerAlaAapArgLeuGlnAapSerValLeuValArgSerPhe 400
DB 1729 AGTGGCGTCAGCTGCGCGACAGCGCTGGAAGACGCAATAGCTCAAGCGCTTC 1788
QY 401 SerLeuAapLeuSerValSerValSerValSerValSerValSerValSerValSer 420
DB 1789 TCTCTGATATCAAAATCAATTTCAATTTCAAGCGATGCGACATCTTCACTGCTTC 1848
QY 421 SerSerSerGluAapAlaLeuGlnArgValArgValArgValArgValArgValArg 440
DB 1849 TCTCTCATCAGAGATGCTTTGGAATCTCAAAACCTTCCACTGCTGATGGGACCAAC 1908
QY 441 LysLeuGlnArgGlnPheSerProValGlnLeuSerGlnGlnArgProGlnArgPro 460
DB 1909 AACCTATGCGAGTCTCCCTGTTCAAGGAATATCGAGACGACTCCGGAACAGCTCT 1968
QY 461 AspLysGlnGluAlaSerValProLysLysLeuGlnThrAlaArgProSerAapSerGln 480
DB 1969 GATTAAGAGGAGACCGACATCCCGACAGACCTGACAGCCCGCTTACAGACGCCAG 2028
QY 481 SerLysArgLeuHisSerValArgThrSerSerSerValThrAlaGlnArgSerLeu 500
DB 2029 AGCAGAGATGATTCGATCGTCAGACACGACGATGCGACCGCCGACAGAGTCCCTTTA 2088
QY 501 SerProLeuHisArgSerGlySerValGlnAapSerValHisThrSerPheLeuPheGly 520
DB 2089 TCTCCACTGATCGAAGTGGAGCGTGGAGACATTAACACACAGCTTCTTTCGCGC 2148
QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlnLeuLysGlyTPHs 540
DB 2149 CTTTCCACCGACGACGACCTTCCAGAGTCTGCTGCGCTTAAAGGCTGGCAC 2208
QY 541 SerAsp1LeuAlaProGlnThrSerThrProSerLeuThrSerSerValArgPheAla 560
DB 2209 TCGGATATCTGGCCCCCGACACTCTACCCCTTCCCTGACGACGCTGTATTTGCC 2268
QY 561 ThrGlnSerSerHisPheValSerAlaSerAlaLeuGlyLysSerAlaSerValSer 580
DB 2269 ACAGAGCTCTCACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2328
QY 581 AlaThrSerSerSerGlnLeuProThrCysGlyAapGlnValLysSerValArgArgArg 600
DB 2329 GCCTACAGCTGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388
QY 601 GlnLysProSerAapArgAlaAapSerArgArgSerValArgLysLeuProPheGln 620
DB 2389 CAAAGCGCAAGTACAGAGCTGACTCGCGCGAGCGTGAATGAAGAGACCCCTTTGAA 2448
QY 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyLysSerValMetSerGluA 640
DB 2449 AACGAGCTTTAAAGCGAGAACCTGCAATGAAATTTGGAAGAGCATCATGATGAGAAC 2508
QY 641 ArgSerArgGlnGlnLeuGlnLysValGlySerGlnSerSerPheSerGlySerMetGln 660
DB 2509 AGGTCACGGGAAAGCTGGGGAAAGTGGGAGTCAAGTCTTTCGGGCAACATGAA 2568
QY 661 IleIleGlnValSer 665
DB 2569 ATCATTAGAGTCTCC 2583

```

RESULT 12
 ABV21080
 ID ABV21080 standard; cDNA; 5145 BP.
 XX
 AC ABV21080;

```

XX 13-SEP-2002 (first entry)
DT Human prostate expression marker cDNA 21071.
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
OS Homo sapiens.
XX MO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001, 2001MO-US005171.
XX 17-FEB-2000; 2000US-0183119P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0253281P.
XX (MILP-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI, 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 3481; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (c) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (d) selecting a composition for inhibiting prostate cancer in a patient;
XX (e) assessing the prostate cell carcinogenic potential of a compound; (f)
XX determining whether prostate cancer has metastasized in a patient; (g)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX SO Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
XX
XX Alignment Scores:
XX Pred. No.: 2.03e-234 Length: 5145
XX Score: 3406.00 Matches: 663
XX Percent Similarity: 99.85% Conservative: 1
XX Best Local Similarity: 99.70% Mismatches: 1
XX Query Match: 99.65% Indels: 0
XX DB: 5 Gaps: 0
XX
XX US-10-029-345A-109 (1-665) x ABV21080 (1-5145)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
DB 589 ATGGCCCATGAGATGATGATGGAATCTCAAAATTTGTTACTGAGAGGTTGGCTGCTGGA 648
QY 21 SerGlyThrGlnLysValLeuLeuIleAapSerArgProPheValGlnArgValSer 40
DB 649 AGTGAACCGAATAAAGTCTGCTAATTTGATGACCGCGCATTTGTGGAATACATATCC 708
QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGlnGln 60
DB 709 CACATTTTGAAGCCATTATATCACTCTCCAACTTATGAAGCAAGGTTGCAACAG 768
QY 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysValHisValAapIleAap 80

```

```

Db      769 GACAAAGCTTATTATACAGAGCTCATCCAGCATTCACAGAAACATAGAGTTGACATTGAT 828
Qy      81 CysSerGlnLysValValValValYrAspGlnSerSerGlnAspValAserLeuSerSer 100
Db      829 TCGAGTCGAGAGGTTGATGTTTACGATCAAGATCCCAAGATGTTCTCTCTCTCTCA 888
Qy      101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu 120
Db      889 GACTGTTTCTCACTGACTTCTGCGGAAACCTGGAAAGAGCTTCACTCTGTTCACTG 948
Qy      121 LeuAlaGlyLysPheAlaGlnPheSerArgCysPheProGlyLeuCysGlyLysSer 140
Db      949 CTGCAAGTGGGTTTGTGAGATTCTCTCGTTGTTTCCCTGCGCTCTGGAAGAAATCC 1008
Qy      141 ThrLeuValProThrCysLieserGlnProCysLeuProValAlaAsnLiesGlyProThr 160
Db      1009 ACTCTAGTCCCTACCTGCACTTCTGCAAGCTTCTTACCTGTTGCCAACATTGGGCCAAC 1068
Qy      161 ArgLysLeuProAsnLeuLysLeuGlyCysGlnArgAspValLeuAsnLysGlyLeu 180
Db      1069 CGAATTCCTCCCAATCTTATCTTGCTGCCAGCAAGATGCTCTCAAGAGAGCTGATG 1128
Qy      181 GlnGlnAsnGlyLiesGlyTrpValLeuAsnLieserTrpCysProLysProAspPhe 200
Db      1129 CAGCAAGATGGGATTCGTTATGTTAATGCCAGCAATACCTGTCCAAGCTGACTTT 1188
Qy      201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlyLysLeu 220
Db      1189 ATCCCGAGTCTCATTTCTGCGTGTCTGTGAATACAGCTTTGTGAAATTTTG 1248
Qy      221 ProTrpLeuAspLysSerValAspPheLiesGlyLysAlaLysLieserAsnGlyCysVal 240
Db      1249 CCGTGTGTCGCAATGAGTATGATTTCATTGAGAAAGCAAAACCTCCCATGGATGTGT 1308
Qy      241 LeuValHisCysLeuAlaGlyLieserArgSerAlaThrLiesAlaLiesAlaTrpLiesMet 260
Db      1309 CTAGTGCATGTTTAGCTGGAGATCTCCGCTCCGCCACATGCTATCCCTACATCATG 1368
Qy      261 LysArgMetAspMetSerLeuAspGlnAlaTrpArgPheValLysGlyLysArgProThr 280
Db      1369 AAGAGATGACATGCTTCTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1428
Qy      281 LIESerProAsnAspPheAsnLeuGlyLiesLeuAspTrpGlyLysLysLiesLysAsn 300
Db      1429 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGCACTATGAGAAAGAAATTAAGAAC 1488
Qy      301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLiesLeuHisLiesGlyLysProAsn 320
Db      1489 CAGACTGGAGCATCAGAGGCCAAAGCAAACTCAAGCTGTGCACTGGAGAGCCAAAT 1548
Qy      321 GlnProValProAlaValSerGlnGlyLysGlnLysSerGlnThrProLeuSerProPro 340
Db      1549 GAACCTGTCCTGCTGTCTCAGAGGGTGGACAGAAAGAGAGAGCCCTCACTCCACACC 1608
Qy      341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyLysGlnArgProValHisProLiesVal 360
Db      1609 TGTGCCGACTCTGCTACTCAGAGGCAAGCAAGCAAAAGCCCGTGCATCCCGCACCGTG 1668
Qy      361 ProSerValProSerValGlnProSerLeuGlnAspSerProLeuValGlnAlaLeu 380
Db      1669 CCCAGCGGCCACAGCGAGCGCTGCTTGAAGAGAGCCCGCTGGTAAAGGCGCTC 1728
Qy      381 SerGlyLeuHisLiesSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 400
Db      1729 AGTGGCTGCACTGTCGAGAGAGAGCTGGAGAGACAGCAATAGCTCAAGGCTTCTTC 1788
Qy      401 SerLeuAspLiesLysSerValSerTrpSerAlaSerMetAlaLiesLeuHisLiesLysPhe 420
Db      1789 TCTCTGATATCAATCATATTCATATTCAGCCAGATGGAGATCTTACATAGGCTTC 1848
Qy      421 SerSerSerGlnAspAlaLeuGlyLysTrpLysProSerThrThrLeuAspGlyThrAsn 440

```

```

Db      1849 TCCCTACAGAGAGATGCTTTGGAAATCTAACAACTTCCACTACTGTGATGGAGCAAC 1908
Qy      441 LysLeuCysGlnPheSerProValGlnGlyLeuSerGlnGlnThrProGlyLysSerPro 460
Db      1909 AACGTATGCAATGCTTCTCCCTGTTTCAGAGAACTATCGAGAGAGACTCCCGAAACAGCTCT 1968
Qy      461 AspLysGlnGlnAlaSerLiesProLysLysLeuGlnThrLiesLiesLiesLiesLies 480
Db      1969 GATTAAGAGAGAGAGCCAGCATCCCAAGAGCTGCAAGCCCGAGGCTTTCAGACAGCAG 2028
Qy      481 SerLysArgLeuHisLiesSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
Db      2029 AGCAAGCAATGTCATTCGCTGAGAACCAAGAGAGAGTGGACCCCGAGAGGCTTCTTTTA 2088
Qy      501 SerProLeuHisArgSerGlySerValGlnAspAsnTrpHisPheThrSerPheLeuPheGly 520
Db      2089 TCTCCACTGCACTCAAGTGGAGGCTGGAGAGACAAATTCACACAGCTTCTTTTCGCGC 2148
Qy      521 LeuSerThrSerGlnGlnHisLiesThrLysSerAlaGlyLeuGlyLysLeuLysGlyTrpHis 540
Db      2149 CTTTCCACAGCCAGCAAGCACTCAGAAAGTGTGCTGCGCTTGAAGGCTGGCAC 2208
Qy      541 SerAspLiesLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTrpPheAla 560
Db      2209 TCGGATATCTTGGCCCCCGAGACTTACCTCTTCCCTGACCAAGAGCTGGATTTTGGCC 2268
Qy      561 ThrGlnSerSerHisPheTrpSerAlaLiesAlaLiesTrpGlyLysSerAlaSerTrpSer 580
Db      2269 ACAGAGTCTCACACTTACTCTCTCTGCTGACCATCTCAGAGAGGAGTGCAGTTACTCT 2328
Qy      581 AlaTrpSerCysSerGlnLeuProThrCysGlyAspGlnValTrpSerValArgArgArg 600
Db      2329 GCCTACAGCTGCAGCAGAGCTGCCACTTGGCAAGCAAGTCTTATTTGTGCGAGCGG 2388
Qy      601 GlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGlnGlnSerProPheGln 620
Db      2389 CAGAAAGCAAGTGAAGAGCTGACTCGCGGCGAGCTGCATGAAGAGACCCCTTTGAA 2448
Qy      621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyLysLeuLiesLiesSerGlnAsn 640
Db      2449 AAGCAGTTTAAACGAGAGAGCTGCCAATGGATTTGGAAGAGCATGTCAGAGAAAC 2508
Qy      641 ArgSerArgGlnGlnLeuGlnGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
Db      2509 AGCTACAGGAGAGAGCTGGGAGAAAGTGGGAGTCAAGTCTTTCGGGACGATGAA 2568
Qy      661 IleLiesGlnValSer 665
Db      2569 ATCATTTGAGGCTCTC 2583

```

RESULT 13
 ABV26680 ID ABV26680 standard; cDNA, 5145 BP.
 AC ABV26680;
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 26671.
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.

	CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
	CC	specification or its complement. (1) is useful for: (a) assessing whether
	CC	a patient is afflicted with prostate cancer; (b) monitoring the efficacy
	CC	progression of prostate cancer in a patient; (c) assessing the efficacy
	CC	of a test compound to inhibit prostate cancer in a patient; (d) assessing
	CC	the efficacy of a therapy for inhibiting prostate cancer in a patient; (e)
	CC	selecting a composition for inhibiting prostate cancer in a patient;
	CC	(f) assessing the prostate cell carcinogenic potential of a compound; (g)
	CC	determining whether prostate cancer has metastasized in a patient; (h)
	CC	assessing the aggressiveness or indolence of prostate cancer in a patient
	CC	; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
	XX	
	SQ	Sequence 5145 BP, 1408 A, 1135 C, 1253 G, 1346 T, 0 U, 3 Other;
	Alignment Scores:	
	Pred. No.:	2.03e-234
	Score:	3406.00
	Percent Similarity:	99.85%
	Best Local Similarity:	99.70%
	Query Match:	99.65%
	DB:	5
	US-10-029-345A-109 (1-665) x ABV20978 (1-5145)	Gaps: 0
QY	1	MetLahisgUwecIlegIYThrgInIleValThrGluArgLeuValAlaleuLeuGlu
Db	589	ATGCGCCATGAGATGATGGACCTCAATGTTACTGAGCGTGTGGCTCTGCTGGAA
QY	21	SerGIYThrgIuysValIleuLeuIleAspSerArgProPheValGluYrAsnThsSer
Db	649	AGTGAACGAAABAGTGCCTGTAATTGATAGCGCGCATTTGTGGAAATCAATCATCC
QY	41	HisIleLeuGluAlaIleasnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn
Db	709	CACATTTTGGAAACCATTAATATCACTGCTCCAAAGCTTATGAACGAAGGTTGCAACNG
QY	61	AspLysValIleuIleThrgIuLeuIleGlnHisSerAlaYrHisIlyLeuValAspIleAsp
Db	769	GACAAAGTGTAAATTCAGAGCTCATCCAGCAATTAAGGAAACATTAAGGTTGACATTCAT
QY	81	CysSerGlnLysValValValIYrAspGlnSerSerGlnAspValAlaSerLeuSerSer
Db	829	TGAGTCAGAAAGGTTGATGATTACATCAAGCTCCAAAGATGTTGGCTCTCTCTTCA
QY	101	AspCysPheLeuThrValIleuLeuGlyIysLeuGluIysSerPheAsnSerValHisLeu
Db	889	GACTGTTTTCACGTACTCTTGGGTAACTGGAGAAAGACTTCACTCTGTTCACTCG
QY	121	LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer
Db	949	CTTGCAGAGGAGGTTGCTGAGTTCCTCTCGTTGTTCCCTGGCTCTGTGAAGGAAATCC
QY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr
Db	1009	ACTCTAGCTCCCTACCTGATTTCTCACGCTTCTTACCTGTGCGCAACATTTGGCCAAACC
QY	161	ArgIleLeuProAsnLeuIYrLeuGlyCysGlnArgAspValLeuAsnIysGluLeuIle
Db	1069	CGAATTTCTCCCAATCTTATCTTGGCTGCACAGCAAGATCTCTCAACAAGAGCTGATG
QY	181	GlnGlnAsnGlyIleGlyIYrValIleAsnAlaSerTyrThrCysProIYrProAspPhe
Db	1129	CAGCAGATAGGATGTTATGTGTTTAATGCCAGCAATTAACCTGTCTCAAAAGCTTACATT
QY	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIYrIleLeu
Db	1189	ATCCCGCAGTCTATTTCTCGCTGCGCTGCGATGACAGCTTTTGTGAGAAATTTTG
QY	221	ProTrpLeuAspIYrSerSerValAspPheIleGlyIYrAlaIYrAlaSerAsnIYrCysVal
Db	1249	CCGTGTTGGCAAAATCAGTAAATTCAATTGGAAGGAAAGGCTCCAAATGAGATGTGT
QY	241	IleValHisCysLeuAsnIleGlyIYrAlaIYrAlaSerAsnIYrCysVal

```

Db      1309 CTAGTGACGTTTAAAGTGGAGTCTCCGCTCCGACCACTCGGTATCGCTCATCATG 1368
Qy      261  LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleuGlyIleuSerGluAsn 280
Db      1369 AAGAGGTGAGCACTGCTTTAGTGAAGCTTACAGATTTTGTGAAGAAAAAGAACCTTACT 1428
Qy      281  IleSerProAsnMetPheLeuGlyIleuLeuAspTyrGlyIleuValIleuAsn 300
Db      1429 ATATCTCCAACTTCAATTTTCTGGCCAACTCTGAGCATATGAGAAAGATTAAAGAAC 1488
Qy      301  GlnThrGlyAlaSerGlyProLysSerLeuLeuIleuLeuIleuGlyIleuProAsn 320
Db      1489 CAGACTGGAGCACTCAGGCGCAAGAGCAAACTCAAGCTGTGCACTCGAGAAAGCCAAAT 1548
Qy      321  GluProValProAlaValSerGlyGlyIleuLysSerGluThrProLeuSerProPro 340
Db      1549 GAACCTGTCCCTCTGTCTCAGAGGCTGAGCAAGAAAGCAGACGCCCTCAGTCCACCC 1608
Qy      341  CysAlaAspSerLeuThrSerGluAlaGlyIleuArgProValIleuAspProAlaSerVal 360
Db      1609 TGTGCCGACTCTCTACTCAGAGGCGAGAGCAAAAGCCCTGTCATCCGCGCAGCGTG 1668
Qy      361  ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
Db      1669 CCAGCGCTGCCCGCGCGAGCGCGCTGTTAGAGACAGCCCGCTGTTACAGCGCTC 1728
Qy      381  SerGlyLeuIleuIleuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
Db      1729 AGTGGCGTGCACCTGTCCGAGAGAGCTGAAAGACGAAATAGCTCAAGCGTTCCTTC 1788
Qy      401  SerLeuAspIleuLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuIleuGlyPhe 420
Db      1789 TCTCTGATATCAATTCATGTTTCAATTCAGCCAGACATGACACATCTTACATGGCTTC 1848
Qy      421  SerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
Db      1849 TCTCTCATCAAAATGCTTTGGAATACTCAAACTTCCACTCTCGATGAGGACCAAC 1908
Qy      441  LysLeuGlyGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
Db      1909 AAGCTATGCCAGTCTCCCTGTTACAGGAATCTGAGAGACAGATCCCGAAACAGAGCTC 1968
Qy      461  AspLysGlnGluAlaSerIleProLysLeuGlnThrAlaArgProSerAspSerGln 480
Db      1969 GATPAGAGGAGAACCCAGCATCCCAAGAACTGAGCCGCGCAGCTTACAGACAGCAG 2028
Qy      481  SerLysArgLeuIleSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db      2029 AGCAGAGCATTCATTCGTCAGAACAGACAGCATGAGCAGCCCGCCAGAGTCCCTTTTA 2088
Qy      501  SerProLeuIleAspSerGlySerValGluAspAsnTyrIleThrSerPheLeuPheGly 520
Db      2089 TCTCCACTGCATCGAAGTGGAGGTGAGAGCAATTACACACAGCTTCCTTTTGGC 2148
Qy      521  LeuSerThrSerGlnGlnIleuThrLysSerAlaGlyLeuGlyIleuLysGlyTyrPheIle 540
Db      2149 CTTTCCACCGACGACGACCTTCAACAACTGCTGCGCTGCGCTTAAAGGCTGCGCAC 2208
Qy      541  SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPyrPheAla 560
Db      2209 TCGGATATCTTGGCCCCCGACACTTCAACCCCTTCCCTGACACAGACCTGATTTTGGC 2268
Qy      561  ThrGluSerSerIlePheTyrSerAlaSerAlaIleTyrGlyIleSerAlaSerTyrSer 580
Db      2269 ACAGAGTCTCACACTTCTACTGCTGCTCAGCCATCTACGAGGAGGAGTGCAGTTACTCT 2328
Qy      581  AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValaArgArg 600
Db      2329 GCGTACAGCTGACGACGCTGCGCACTGCGAGAACCAACTCTAATCTTGCGCAGCGCG 2388
Qy      601  GlnLysProSerAspArgAlaAspSerArgSerTyrIleGluIleProPheGlu 620

```

```

Db      2389 CAGAACCAAGTACAGAGCTGACTCGCGCGAGCTGGCATGAAAGAGCCCCTTTGA 2448
Qy      621  LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleuMetSerGluAsn 640
Db      2449 AAGCAGTTTAAAGCAGAAAGCTGCGCAATGGAATTTGGAGAGGACATCACTCAGAGAAC 2508
Qy      641  ArgSerArgGluIleuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db      2509 AGGTCAAGGAGAGAGCTGGGGAAGTGGCAGTCACTTACCTTTTCCGCGACATGAA 2568
Qy      661  IleIleGluValSer 665
Db      2569 ATCATTGAGGTCTCC 2583

RESULT 15
ABV21092
ID      ABV21092 standard; cDNA; 5145 BP.
XX      XX
AC      ABV21092;
XX      XX
DT      13-SEP-2002 (first entry)
XX      XX
DE      Human prostate expression marker cDNA 21093.
XX      XX
KW      Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
        pharmacogenomic marker; gene; ss.
XX      XX
OS      Homo sapiens.
XX      XX
PN      MO200106060-A2.
XX      XX
PD      23-AUG-2001.
XX      XX
PF      20-FEB-2001; 2001MO-US005171.
XX      XX
PR      17-FEB-2000; 2000US-0183319P.
PR      16-MAR-2000; 2000US-0189862P.
PR      25-MAY-2000; 2000US-0207454P.
PR      09-JUN-2000; 2000US-0211314P.
PR      18-JUL-2000; 2000US-0219007P.
PR      13-DEC-2000; 2000US-0255281P.
XX      XX
PA      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX      XX
PI      Schlegel R, Endege WO, Monahan JR;
XX      XX
DR      WPI; 2001-662795/76.
XX      XX
PT      Novel isolated nucleic acid molecule associated with cancerous state of
        prostate cells and correlating with presence of prostate cancer, useful
        for detecting presence of prostate cancer, stage of prostate cancer.
XX      XX
PS      Claim 1; Page 3485; 11750bp; English.
XX      XX
CC      The invention relates to an isolated nucleic acid molecule (i) comprising
        a nucleotide sequence given in Tables 1-9 (ABY00010-ABY62213) of the
        specification or its complement; (ii) is useful for: (a) assessing whether
        a patient is afflicted with prostate cancer; (b) monitoring the
        progression of prostate cancer in a patient; (c) assessing the efficacy
        of a test compound to inhibit prostate cancer in a patient; (d) assessing
        the efficacy of a therapy for inhibiting prostate cancer in a patient;
        (e) selecting a composition for inhibiting prostate cancer in a patient;
        (f) assessing whether prostate cancer has metastasized in a patient; (g)
        determining whether prostate cancer has metastasized in a patient; (h)
        assessing the aggressiveness or indolence of prostate cancer in a patient
        ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
        XX
SQ      Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Alignment Scores:
Pred. No.: 2,03e-234 Length: 5145
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1

```


Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: 5 Gaps: 0

US-10-029-345A-109 (1-665) x ABV21092 (1-5145)

QY 1 MetalaHISGLUmetilegLYThrglnileValThrgluArgleuValaleuLeuGlu 20
 DB ATGGCCCATGAGATATTGGAACCAAAATTGTTACTGAGAGTTGGTGGCTCTGCTGGA 648
 QY 21 SerGlyThrgluValleuLeuileaspSerArgProPheValGluTYrAsnThsSer 40
 DB AGTGGAAGCAAAAGTCTGCTAATTGATAGCCGCCCATTTGGAAATACAAATCAATCC 708
 QY 41 HisileuGluAlaileasnilleasnCyserSerlyseuMetlyseuArgleuGln 60
 DB CACATTGTGAAAGCCATTAAATCAACTGCTCCAAAGCTTAATGAAGCGAAAGTTGCAAG 768
 QY 61 AsplysValleuileThrgluLeuileglnHISerialySHIslyValAspIleasp 80
 DB GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGGAAACAAATAGTTGACATTGAT 828
 QY 81 CysSerGlnlyseValValValTYrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB TGCAGTCAAGAGTTGTGTATTCATCAAGCTCCCAAGATGTTGCCCTCTCTCTTCA 888
 QY 101 AspCysPheLeuThrValleuLeuGlylyseuGlylyseuPheAsnSerValHISleu 120
 DB GACTGTTTTCACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
 QY 121 LeuAlaGlyGlyPheAlaGlyPheSerArgCysPheProGlyLeuGlyGlylyseuSer 140
 DB CTTCAGAGGTGGTGTGCTGAGATTCTCTGTTGTTCTCTGCTGCTGCTGCTGCTGCTG 1008
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 1009 ACTCAAGTCCCTACCTGCAATTCCTCAAGCTTCCTGCTGCTGCTGCTGCTGCTGCT 1068
 QY 161 ArgIleLeuProAsnLeuTYrLeuGlyCysGlnArgAspValleuAsnlyseuIle 180
 DB 1069 CGAATTCCTCCCAATCTTAACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1128
 QY 181 GlnGlnAsnGlyIleGlyTYrValleuAsnAlaSerTYrThrCysProlyAspAspPhe 200
 DB 1129 CAGCAGAAATGGATTGTTATGTTAAATGCAAGCAATACCTGCTCAAAAGCTGACTT 1188
 QY 201 IleProGlnSerHISpHeLeuArgValProValAsnAspPheCysGlylyseuIleu 220
 DB 1189 ATCCCGAGTCTCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
 QY 221 ProTrpLeuAspLyseSerValAspPheIleGlnlyseAlaIleAsnGlyCysVal 240
 DB 1249 CCGTGTGGTGAACAAATGATGATTTCAATTGAAAGCAAAAGCTCCAAATGATGCTT 1308
 QY 241 LeuValHISySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTYrIleu 260
 DB 1309 CTAGGCACTGTTAGTGGGATCTCCCGCTCCGCAACATGCTATGCTTACATCATG 1368
 QY 261 LysArgMetAspMetSerLeuAspGlnAlaTYrArgPheVallyseGlnlyseArgProThr 280
 DB 1369 AAGGAGATGACATGCTTATGATGAACCTTACAGATTGTGAAAGAAAAAGCACTACT 1428
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnleuLeuAspTYrGlnlyseIleIleAsn 300
 DB 1429 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGAGCTATGAAAGAAATTAAGAAC 1488
 QY 301 GlnThrglyAlaSerGlyProlyseSerlyseuLyseuIleuHISleuGlnlyseProAsn 320
 DB 1489 CAGACTGGAGCATGAGGCGCAAAAGCAAACTCAAGCTGCTGCACTGGAGAAACCAAT 1548
 QY 321 GlnProValProAlaValSerGlyGlyGlnlyseSerGlyuThrProLeuSerProPro 340
 DB 1549 GAACCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1608

QY 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHISProAlaSerVal 360
 DB 1609 TGTGCCACTCTGCTACTTACAGAGGACAGAGCAAAAGCCCTGCTGCTCCCGCAGGCTG 1668
 QY 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
 DB 1669 CCGAGCTGCCAGCGTGCAGCGCTGCTGTAGAGGACAGCCGCTGTGATACAGGCGCTC 1728
 QY 381 SerGlyIleuHISleuSerAlaAspArgleuGlnAspSerlyseuLyseuAspSerPhe 400
 DB 1729 AGTGAGCTGACCTGTCTGAGACAGAGCTGAAAGACAAATAGCTCAAGCGTCTCTTC 1788
 QY 401 SerLeuAspIleIlyseSerValSerTYrSerAlaSerMetAlaIleAsnIleGlyPhe 420
 DB 1789 TCTTGATATTCAAATAGATTTTCAATATTCAATTCAGCCAGCATGCGACATCTTACATGCTTC 1848
 QY 421 SerSerSerGlnAspAlaLeuGlnlyseTYrlyseProSerThrThreAspGlyThrAsn 440
 DB 1849 TCTCATCAGAAAGATGCTTTGGAATACTCAAACTTCACACTCTGATGGAGCAAC 1908
 QY 441 LysLeuCySglnPheSerProValGlnIleuSerGlnIleThrProGlnThrSerPro 460
 DB 1909 AAGCTATGCCAGTCTCTCCCTGTCAGAACTATCGAGAGACTCCCGAAACCAATCTT 1968
 QY 461 AsplysGlnAlaSerIleProLylyseuGlnThrAlaArgProSerAspSerGln 480
 DB 1969 GATPAGAGAGAACCCAGCATCCCAAGAACTCCAGCCAGCTTTCAGACAGCAG 2028
 QY 481 SerlyseArgleuHISerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 DB 2029 AGCAAGGATTTGATTCGTGCAGAACCAAGCAGAGTGGCACCCCAAGAGTCCCTTTTA 2088
 QY 501 SerProLeuHISArgSerGlySerValGlnAspAsnTYrHISerSerPheLeuPheGly 520
 DB 2089 TCTCACTGCATGAAAGTGGAGGTGGAGGAGAAATTAACCAACCAAGCTTCTTTTCGCT 2148
 QY 521 LeuSerThrSerGlnGlnHISleuThrlyseSerAlaGlyLeuGlylyseuGlyTYrPhe 540
 DB 2149 CTTTCCACCAAGCCAGACGACACTCAAGAACTGCTGCGCTGGCCCTTAAGGCTGGCAC 2208
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTYrPheAla 560
 DB 2209 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGGTATTTTGGCC 2268
 QY 561 ThrGlnSerSerHISpHeTYrSerAlaSerAlaIleTYrGlylyseSerAlaSerTYrSer 580
 DB 2269 ACGAGTCTCAACAATTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2328
 QY 581 AlaTYrSerCysSerGlnleuProThrCysGlyAspGlnValTYrSerValArgArgArg 600
 DB 2329 GCTTACAGCTGACGACGACGCTGCCACTTGGCGAGACCAAGTCTAATCTTGCGCAGGCGG 2388
 QY 601 GlnlyseProSerAspArgAlaAspSerArgArgSerTYrPheGlnlyseProPheGln 620
 DB 2389 CAGAAGCAAGTACAGAGCTGACTCGCGCGAGCTGGCATTAAGAGAGGCCCTTTTGA 2448
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnIleuSerIleMetSerGlnAsn 640
 DB 2449 AAGCAGTTTAAAGCAGAACTCCCAATGGAATTTGGAGAGGCATCATGTCAAGAAC 2508
 QY 641 ArgSerArgGlnIleuGlylyseValGlySerGlnSerSerPheSerGlySerMetGln 660
 DB 2509 AGGTCAAGGAAAGCTGGGGAAGTGGGAGTCAAGTCTTCTTCCGCAACATGGA 2568
 QY 661 IleIleGlnValSer 665
 DB 2569 ATCATTGAGGCTTCC 2583

Search completed: June 21, 2004, 21:32:57
 Job time : 776.933 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:10:27 ; Search time 74.2709 Seconds

(without alignments)
2529.847 Million cell updates/sec

Title: US-10-029-345A-109

Sequence: 1 MAHEMIGQIVTERLVALL.....LGRVSGSSFSGSMETIEVS 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.29Jan04:*

1: geneseqp1980s:***
2: geneseqp1980s:***
3: geneseqp2000s:***
4: geneseqp2001s:***
5: geneseqp2002s:***
6: geneseqp2003as:***
7: geneseqp2003bs:***
8: geneseqp2004s:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3418	100.0	665	5	ABR52381 Protein r
2	3418	100.0	665	5	ABR52407 Protein r
3	3406	99.6	665	4	AAE04834 Human SGP
4	3406	99.6	665	4	AAU09016 Human dua
5	3406	99.6	665	5	AAU79156 Human dua
6	3406	99.6	665	5	AAU09946 Protein s
7	3406	99.6	665	5	AAU75789 Human pro
8	3406	99.6	665	5	ABR97946 Human pro
9	3406	99.6	665	5	AAU79929 Human dua
10	3406	99.6	665	5	ABR97291 Novel hum
11	3399	99.4	665	6	ADA54744 Human pro
12	3398	99.4	665	5	AAU79161 Human dua
13	3396	99.4	665	5	AAU79162 Human dua
14	3388	99.1	665	5	ABR52352 Protein r
15	3368.5	98.6	664	5	ABR52424 Protein r
16	3057.5	89.5	660	5	ABR52385 Protein r
17	3030.5	88.7	666	4	AAE03325 Human pro
18	2930	85.7	672	4	AAW25744 Human pro
19	2500	73.1	517	5	AAU79159 Human dua
20	1542	45.1	302	5	ABR52425 Protein r
21	1326	38.8	625	5	ABR52382 Protein r
22	1326	38.8	625	5	ABR52350 Protein r
23	1326	38.8	625	6	ABG73440 Human dua
24	1302	38.1	663	5	ABR52351 Protein r
25	1297	37.9	663	2	AAW29150 Dual-spec

26	977.5	28.6	579	7	ABR08458	Ado08458 Novel pro
27	742.5	21.7	375	4	ABG07902	Abg07902 Novel hum
28	741	21.7	140	5	ABR52404	AbR52404 Peptide r
29	716	20.9	140	5	ABR52405	AbR52405 Peptide r
30	654.5	19.1	170	4	ABR6436	AbR6436 Human MAP
31	654.5	19.1	170	4	AAE06780	AAE06780 Human dua
32	476	13.9	482	4	ABR52325	AbR52325 Human pro
33	471.5	13.8	381	5	ABR52383	AbR52383 Protein r
34	471.5	13.8	381	6	ABR66803	AbR66803 Human COP
35	469	13.7	444	5	AAO20515	AAO20515 Protein o
36	469	13.7	482	3	AAI18655	AAI18655 A human r
37	469	13.7	482	3	AAE29641	AAE29641 Human dua
38	469	13.7	482	5	ABR52384	AbR52384 Protein r
39	469	13.7	482	5	ABJ05600	AbJ05600 Breast ca
40	466.5	13.6	381	6	ABP57087	ABP57087 Mouse MKP
41	465.5	13.6	381	7	ADD48300	Add48300 Rat Prote
42	465.5	13.6	381	7	ADG62625	AdG62625 Rat Prote
43	453	13.3	394	4	AAE67875	AAE67875 Human lun
44	453	13.3	394	5	AAU85530	AAU85530 STY8 lung
45	453	13.3	394	6	ABU69502	ABU69502 Humna STY

ALIGNMENTS

RESULT 1	ABR52381	ABR52381 standard; protein; 665 AA.
XX	AC	ABR52381;
XX	AC	19-JUN-2003 (first entry)
DT	XX	Protein relating to the invention SEQ ID NO: 109.
XX	XX	anti-proliferative; hepatotropic; nephrotropic; antiarthritic;
XX	KW	antiproliferative; cardiatic; cytosolic; gene therapy; liver disease;
KW	KW	proliferative disorder; renal failure; cardiovascular disorder;
KW	KW	immunological disorder; arthritis; psoriasis; congenital heart defect;
XX	KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
OS	XX	Homo sapiens.
XX	PN	WO200257460-A2.
XX	PD	25-JUL-2002.
XX	PF	20-DEC-2001; 2001WO-US050459.
XX	PR	20-DEC-2000; 2000US-0256868P.
PR	PR	30-MAR-2001; 2001US-0280186P.
PR	PR	01-MAY-2001; 2001US-0287735P.
PR	PR	05-JUN-2001; 2001US-0295848P.
PR	PR	25-JUN-2001; 2001US-0300465P.
XX	XX	(BRIM) BRISTOL-MYERS SQUIBB CO.
PA	XX	Jackson DG, Feder J, Nelson T, Mintier G, Ramamathan C, Lee L;
XX	PI	Stemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
PI	PI	Kyatek S, Mcatee P, Suchard S, Banas D;
XX	XX	WPI: 2002-599721/64.
DR	XX	N-PSDB; ACC0559.
XX	XX	Novel polynucleotides encoding human phosphatase polypeptides useful in
PT	PT	the prevention or treatment of e.g. proliferative and cardiovascular
PT	PT	disorders.
XX	XX	Claim 5; Fig 12; 801p; English.
PS	XX	The invention relates to a novel isolated nucleic acid comprising a
XX	XX	polynucleotide having a nucleotide sequence selected from 40
CC	CC	polynucleotides fully defined in the specification. The polynucleotide of

CC the invention has antiproliferative, hepatotropic, nephrotropic,
 CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 XX

SO Sequence 665 AA;

Query Match 100.0%; Score 3418; DB 5; Length 665;
 Best Local Similarity 100.0%; Pred. No. 3.1e-257;
 Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAHEMIGIYTERLVALLESGTEKVLIDSRPVEVNTSHLEAININSGKLMKRLQ 60
 DB 1 MAHEMIGIYTERLVALLESGTEKVLIDSRPVEVNTSHLEAININSGKLMKRLQ 60
 QY 61 DKVLITELIHSAAKHKVIDCSQKVVVYDSSQDVASLSDDCELVYLGLKLEKSFNSVHL 120
 DB 61 DKVLITELIHSAAKHKVIDCSQKVVVYDSSQDVASLSDDCELVYLGLKLEKSFNSVHL 120
 QY 121 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGPRLIPNLVILGORDVANKELI 180
 DB 121 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGPRLIPNLVILGORDVANKELI 180
 QY 181 QONGIGVYLNASYCPCPDPIPESHFLRPVNSFCEKILPWLDSKVDPIERAKASNGCV 240
 DB 181 QONGIGVYLNASYCPCPDPIPESHFLRPVNSFCEKILPWLDSKVDPIERAKASNGCV 240
 QY 241 LVHCLAGISRSATTAIYIMKRMDSLDEAYRPEKREPTLSPNPFILGQLIDYEKKIKN 300
 DB 241 LVHCLAGISRSATTAIYIMKRMDSLDEAYRPEKREPTLSPNPFILGQLIDYEKKIKN 300
 QY 301 QTGASGPKSLKLLHLEKNEPVPAYSEGGSSETPPLSPCADSATSTSAAGQRPVHPASV 360
 DB 301 QTGASGPKSLKLLHLEKNEPVPAYSEGGSSETPPLSPCADSATSTSAAGQRPVHPASV 360
 QY 361 PSVPSVQPSLLEDSPLVQALSGHLHSADELEDSNKLKRSFSLDIKSVYSASMAASLHGF 420
 DB 361 PSVPSVQPSLLEDSPLVQALSGHLHSADELEDSNKLKRSFSLDIKSVYSASMAASLHGF 420
 QY 421 SSSEDALEYKRPSTLDTGNKLCQSPVQELSEOTPEPSPKKEASIPKILQIARPSDQ 480
 DB 421 SSSEDALEYKRPSTLDTGNKLCQSPVQELSEOTPEPSPKKEASIPKILQIARPSDQ 480
 QY 481 SKRLHSVTSSSGTAQRSLPLHRSQSVEDNVTSTFLGLSTSQOHLTKSAGLGLKGMH 540
 DB 481 SKRLHSVTSSSGTAQRSLPLHRSQSVEDNVTSTFLGLSTSQOHLTKSAGLGLKGMH 540
 QY 541 SDILAPQSTPSLTSSWTFATESHFPASAIYGGASAYAGSCQLPTCGQVYSVRAR 600
 DB 541 SDILAPQSTPSLTSSWTFATESHFPASAIYGGASAYAGSCQLPTCGQVYSVRAR 600
 QY 601 QKPSDRADSRSMHEEPFEKOFRRSCQWEPGESIMSENRSREELGKYGSSSFGSME 660
 DB 601 QKPSDRADSRSMHEEPFEKOFRRSCQWEPGESIMSENRSREELGKYGSSSFGSME 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 2
 ID ABR52407 standard; protein; 665 AA.
 XX ABR52407;
 XX

DT 19-JUN-2003 (first entry)

XX Protein relating to the invention SEQ ID NO: 148.

DE antiproliferative; hepatotropic; nephrotropic; antiarthritic;
 XX antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
 KW proliferative disorder; renal failure; cardiovascular disorder;
 KW immunological disorder; arthritis; psoriasis; congenital heart defect;
 XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
 OS Homo sapiens.
 XX

PN WO200257460-A2.

PD 25-JUL-2002.

PF 20-DEC-2001; 2001WO-US050459.

PR 20-DEC-2000; 2000US-0256868P.

PR 30-MAR-2001; 2001US-0280186P.

PR 01-MAY-2001; 2001US-0287735P.

PR 05-JUN-2001; 2001US-0295648P.

PR 25-JUN-2001; 2001US-0300465P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
 PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
 PI Krystek S, Mcatee P, Suchard S, Banas D;
 XX WPI; 2002-599721/64.
 DR N-PSDB; ACC60572.

PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.

XX Disclosure; Fig 19; 801pp; English.

CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatotropic, nephrotropic,
 CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 XX

SO Sequence 665 AA;

Query Match 100.0%; Score 3418; DB 5; Length 665;
 Best Local Similarity 100.0%; Pred. No. 3.1e-257;
 Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAHEMIGIYTERLVALLESGTEKVLIDSRPVEVNTSHLEAININSGKLMKRLQ 60
 DB 1 MAHEMIGIYTERLVALLESGTEKVLIDSRPVEVNTSHLEAININSGKLMKRLQ 60
 QY 61 DKVLITELIHSAAKHKVIDCSQKVVVYDSSQDVASLSDDCELVYLGLKLEKSFNSVHL 120
 DB 61 DKVLITELIHSAAKHKVIDCSQKVVVYDSSQDVASLSDDCELVYLGLKLEKSFNSVHL 120
 QY 121 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGPRLIPNLVILGORDVANKELI 180
 DB 121 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGPRLIPNLVILGORDVANKELI 180
 QY 181 QONGIGVYLNASYCPCPDPIPESHFLRPVNSFCEKILPWLDSKVDPIERAKASNGCV 240

```

Db      181 QQNGIGVNLASATCPKDPDPESHFLRVVNSFCCKILPMDLKSVDPIEKKASNGCV 240
Qy      241 LVHCLAGISRSATTAIAYIMKRDMSIDEAYRVEKERPTISPNFNLGOLLDEKKIKN 300
Xx      241 LVHCLAGISRSATTAIAYIMKRDMSIDEAYRVEKERPTISPNFNLGOLLDEKKIKN 300
Qy      301 QTGASGPKSKLKLHLKEPNEPVAVSEGGQKSTPLSPCADSATSEAGRPVHPASV 360
Db      301 QTGASGPKSKLKLHLKEPNEPVAVSEGGQKSTPLSPCADSATSEAGRPVHPASV 360
Qy      361 PSVPSVQPSLLEDSPLVQALSGHLSDRLSDSNKLKRSPLDIKSYSGASMAASLHGF 420
Db      361 PSVPSVQPSLLEDSPLVQALSGHLSDRLSDSNKLKRSPLDIKSYSGASMAASLHGF 420
Qy      421 SSSSDALEYKPKSTTLDGTNKLCOFSPVOELSEQTPEPSPDKEASIPKKLQTPARPSDSQ 480
Db      421 SSSSDALEYKPKSTTLDGTNKLCOFSPVOELSEQTPEPSPDKEASIPKKLQTPARPSDSQ 480
Qy      481 SKRLHSVRTSSGTAORSLSPHRSQSVEDNHTSLFGLSTSQOHLTKSAGLGLKGMH 540
Db      481 SKRLHSVRTSSGTAORSLSPHRSQSVEDNHTSLFGLSTSQOHLTKSAGLGLKGMH 540
Qy      541 SDILAPQTSPTSLTSSWYFATESHFFASAIYQGSASVAVSCQPLTGQDYVSVRRR 600
Db      541 SDILAPQTSPTSLTSSWYFATESHFFASAIYQGSASVAVSCQPLTGQDYVSVRRR 600
Qy      601 QKPSDRADSRSSWHEESPFEKQFRRSCQMEFGESINENSRRELKGVSSQSSFGSME 660
Db      601 QKPSDRADSRSSWHEESPFEKQFRRSCQMEFGESINENSRRELKGVSSQSSFGSME 660
Qy      661 IIEVS 665
Db      661 IIEVS 665

RESULT 3
AAE04834
ID      AAE04834 standard; protein; 665 AA.
XX
AC      AAE04834;
XX
DT      10-SEP-2001 (first entry)
XX
DE      Human SGP002 phosphatase polypeptide.
XX
KW      Human; SGP002 phosphatase polypeptide; phosphatase-related disease;
KW      immune-related disorder; ocular disease; organ transplant rejection;
KW      infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW      metabolic disorder; hematopoietic cancer; mood disorder; cardiac;
KW      Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW      cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW      attention disorder; cognition disorder; psychotic disorder; cytostatic;
KW      neurological disorder; vitruclide; nootropic; cerebroprotective; therapy;
KW      neuroprotective; antibacterial; antineoplastic; tranquilizer; antidiabetic;
KW      hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW      antifungal; dual specificity phosphatase; DSP; Map kinase phosphatase;
KW      MKP; migraine; chromosome 12p11.1-p12.1.
XX
OS      Homo sapiens.
XX
PH      Key      Location/Qualifiers
FT      Domain      1..173
FT      Domain      /label= Catalytic_domain
FT      Domain      158..297
FT      Domain      /label= Phosphatase_domain
XX
PN      MO200146394-A2.
XX
PD      28-JUN-2001.
XX
PF      21-DEC-2000; 2000MO-US034736.
XX

```

```

PR      21-DEC-1999; 99US-0173255P.
PR      28-DEC-1999; 99US-0175766P.
PR      25-JAN-2000; 2000US-0178078P.
PR      31-JAN-2000; 2000US-0179301P.
XX
PA      (SUGB-) SUGEN INC.
XX
PI      Plozman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;
PI      Hill RJ, Flanagan P;
XX
DR      WPI; 2001-418058/44.
DR      N-PSDB; AAD09492.
XX
PT      Novel phosphatase polypeptide useful for treating cancers, immune-related
PT      diseases and disorders, cardiovascular disease, brain or neuronal-
PT      associated diseases and metabolic disorders.
XX
PS      Claim 7; Fig 2; 186pp; English.
XX
CC      The present invention relates to phosphatase polypeptides, nucleotide
CC      sequences encoding them, as well as various products and methods useful
CC      for the diagnosis and treatment of various phosphatase-related diseases
CC      and conditions. Substance that modulates the activity of phosphatase
CC      polypeptide is used to treat immune-related diseases and disorders,
CC      cardiovascular disease, brain or neuronal-associated diseases and
CC      metabolic disorders, including cancers of tissues, cancers of
CC      hematopoietic origin, diseases of central and peripheral nervous system,
CC      Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
CC      lateral sclerosis, viral infections, infections caused by prions, sexual
CC      bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual
CC      dysfunction, mood disorders, attention disorders, cognition disorders,
CC      hypotension, hypertension, psychotic disorders, neurological disorders,
CC      dyskinesias and organ transplant rejection. The present amino acid
CC      sequence is human SGP002 phosphatase polypeptide. This sequence is
CC      classified as dual specificity phosphatase (DSP) and Map kinase
CC      phosphatase (MKP). SGP002 gene maps to chromosomal position 12p11.1-p12.1.
XX
SQ      Sequence 665 AA:
XX
Query Match      99.6%; Score 3406; DB 4; Length 665;
Best Local Similarity 99.7%; Pred. No. 2,7e-256;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy      1 MAHEMGTQIVTERLVALLSEGTKEKYLIDSRFVEVNTSHIEATININSGKMKRRLQ 60
Db      1 MAHEMGTQIVTERLVALLSEGTKEKYLIDSRFVEVNTSHIEATININSGKMKRRLQ 60
Qy      61 DKYLITELQHSAMKHKVDIDCSQKVVYDQSSQDVASLSSDCEFTVLLGLKESFNSVHL 120
Db      61 DKYLITELQHSAMKHKVDIDCSQKVVYDQSSQDVASLSSDCEFTVLLGLKESFNSVHL 120
Qy      121 LAGFAEFRCRCPGLCEGKSTLVPTCISQCLPVANIGPRTIIPNLYLGGQRDVANKEL 180
Db      121 LAGFAEFRCRCPGLCEGKSTLVPTCISQCLPVANIGPRTIIPNLYLGGQRDVANKEL 180
Qy      181 QQNGIGVNLASATCPKDPDPESHFLRVVNSFCCKILPMDLKSVDPIEKKASNGCV 240
Db      181 QQNGIGVNLASATCPKDPDPESHFLRVVNSFCCKILPMDLKSVDPIEKKASNGCV 240
Qy      241 LVHCLAGISRSATTAIAYIMKRDMSIDEAYRVEKERPTISPNFNLGOLLDEKKIKN 300
Db      241 LVHCLAGISRSATTAIAYIMKRDMSIDEAYRVEKERPTISPNFNLGOLLDEKKIKN 300
Qy      301 QTGASGPKSKLKLHLKEPNEPVAVSEGGQKSTPLSPCADSATSEAGRPVHPASV 360
Db      301 QTGASGPKSKLKLHLKEPNEPVAVSEGGQKSTPLSPCADSATSEAGRPVHPASV 360
Qy      361 PSVPSVQPSLLEDSPLVQALSGHLSDRLSDSNKLKRSPLDIKSYSGASMAASLHGF 420
Db      361 PSVPSVQPSLLEDSPLVQALSGHLSDRLSDSNKLKRSPLDIKSYSGASMAASLHGF 420
Qy      421 SSSSDALEYKPKSTTLDGTNKLCOFSPVOELSEQTPEPSPDKEASIPKKLQTPARPSDSQ 480
Db      421 SSSSDALEYKPKSTTLDGTNKLCOFSPVOELSEQTPEPSPDKEASIPKKLQTPARPSDSQ 480

```

CC nucleic acid fragments are useful for detecting the presence of the
CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
CC antisense sequences) are useful for modulating the activity or expression
CC of the polypeptide/nucleic acid, useful for the treatment of e.g.
CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
CC listed in the specification), liver disorders, erythroid associated
CC disorders (e.g. hemolytic anemia), cellular proliferative or
CC differentiative disorders, leukemias (e.g. acute myeloid leukaemia),
CC sarcomas, metastatic cancers. Agents modulating 2117 and
CC 36692 are also useful for modulating the proliferation, survival,
CC migration or differentiation of a 36692 or 2117-expressing cell. The
CC polypeptide and nucleic acids are useful for identifying modulating
CC agents. The present sequence represents the dual specificity phosphatase
CC 2117

Sequence 665 AA;

Query Match	99.6%;	Score 3406;	DB 4;	Length 665;
Best Local Similarity	99.7%;	Pred. No. 2.7e-256;		
Matches 663; Conservative	1;	Mismatches 1;		

1; Mismatches 1; Indels 0; Gaps 0;

1 MAHEMIGTIVTERLVALLSEGTAKLLIDSRPFVEYNTSHILEAININCSKLMKRLQQ 60

61 DKVITELIQSAKHKVDIDCSQKVVVYDQSSQDVASLSSDCFIVLLGKLEKSFNSVHL 12

121 LAGGFAEFSRCFQGLCEGKSTLVPTCISQPLPVANIGFTRILENLYLGGQRDVANKELM 18

181 QONGIGVLNANTCPKPDFIESHFLRPVAVDSFCEKILPWLKDSVDFIEKASNGCV 24

241 LVHCLGISRSATLAIAYIMKMDMSIDEAYFVKEKPTISNENFLGQLDYEEKIKN 300

301 QTSAGSGFASLKLTLLHEKPNHPVPAVSEGGQKSETPLSPPCADSATSEAAAGRPVHPASV 360
301 QTGAGPKSKLTLLHEKPNHPVPAVSEGGQKSETPLSPPCADSATSEAAAGRPVHPASV 360

361 PSVPSVQPSLLEDSPLVQALSGHLHSADRLIEDSNKLKRSFSLDIKVSYSASMAASLHGF 420
361 PSVPSVQPSLLEDSPLVQALSGHLHSADRLIEDSNKLKRSFSLDIKVSYSASMAASLHGF 420

```

421 SSSSEDALRYKKPSTTLDGINKLCOFPSVQELSEQPETSDDKEASIPKLTQTARPSDQ 480
|||||
421 SSSSEDALRYKKPSTTLDGINKLCOFPSVQELSEQPETSDDKEASIPKLTQTARPSDQ 480

```

481 SKRLHVSRTSSGTAQRSLSPHRRSGVEDNHTSFLFGLSTSOQHLTKSAGLCKWH 540
|||||
481 SKRLHVSRTSSGTAQRSLSPHRRSGVEDNHTSFLFGLSTSOQHLTKSAGLCKWH 540
|||||

541 SDILAPQTSTPSLTSSWYFATSSSHFYASASAIYGSASYSQSLPTCCGDQYYSVRRR 600
|||||
541 SDILAPQTSTPSLTSSWYFATSSSHFYASASAIYGSASYSQSLPTCCGDQYYSVRRR 600
|||||

601 QKPSDADSRSSWHEESPFEQKFKRSSQMEGCSIMSENKREELKTVGSQSSFGSGME 660
|||||
601 QKPSDADSRSSWHEESPFEQKFKRSSQMEGCSIMSENKREELKTVGSQSSFGSGME 660
|||||

```

661 IIEVS 665
      ||||
661 IIEVS 665

```

RESULT 5
AAU79156

SULT 4
 U09016
 AAU09016 standard; protein; 665 AA.
 AAU09016;
 18-DEC-2001 (first entry)
 Human dual specificity phosphatase 21117.
 Human; dual specificity phosphatase 21117; hepatotropic; cytostatic;
 hemetopoleic disorder; autoimmune disorder; diabetes mellitus;
 rheumatoid arthritis; multiple sclerosis; Crohn's disease;
 liver disorder; erythroid associated disorder; haemolytic anaemia;
 cellular proliferative, differentiative disorder; leukaemia;
 acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer;
 immunogen.
 Homo sapiens.

Key	Location/Qualifiers
Domain	11..131
Domain	/label= Rhodanese-like domain
Domain	158..297
Active-site	/label= Catalytic domain
	/note= "Dual specificity phosphatase catalytic domain"
	242..254
	/label= Tyrosine-specific_protein_phosphatase_active_site

 MO20017059-A2.
 04-OCT-2001.
 23-MAR-2001; 2001MO-US009477.
 24-MAR-2000; 2000US-0191858P.
 (MILL-) MILLENNIUM PHARM INC.
 Meyers RA;
 WPI; 2001-611635/70.
 N-PsDB; AAS14639.
 New human dual specificity polypeptides and nucleic acids for diagnosis
 of disease and treatment of e.g. liver disorders.
 Claim 9; Fig 1; 143pp; English.
 The invention relates to two novel human dual specificity phosphatases
 designated 21117 and 38692, the nucleic acids encoding them (including
 fragments, allelic variants, their complements or nucleic acids that
 hybridise to them) and antibodies raised against the proteins. The
 antibody is useful for detecting the presence of the polypeptide, and the

ID AAU79156 standard; protein; 665 AA.
 AC AAU79156;
 DT 02-JUL-2002 (first entry)
 DE Human dual-specificity phosphatase-3 (DSP-16) protein.
 XX
 XX Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; chromosome 12p.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT Domain 242..251
 FT /label= Active_site_domain
 XX
 XX MO200226997-A2.
 PD 04-APR-2002.
 PF 25-SEP-2001; 2001MO-US030124.
 XX
 XX 26-SEP-2000; 2000US-0235487P.
 PR
 XX (CEPT-) CEPTYR INC.
 PA
 XX
 PI Luche RM, Wei B;
 XX
 DR WPI; 2002-315802/35.
 DR N-PSDB; ABK47596.
 XX
 XX New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX
 PS Claim 1; Fig 2; 87pp; English.
 XX
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival
 CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC protein of the invention. This sequence is encoded by the human DSP-16
 CC gene located on chromosome 12p
 CC
 CC
 SO Sequence 665 AA;
 Query Match 99.6%; Score 3406; DB 5; Length 665;
 Best Local Similarity 99.7%; Pred. No. 2.7e-256;
 Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MAHEMIGTQVTERLVALLSGTEKVLIDSRPVEVNTSHILEAININSGKMKRRLQ 60
 DB 1 MAHEMIGTQVTERLVALLSGTEKVLIDSRPVEVNTSHILEAININSGKMKRRLQ 60
 QY 61 DKVLITELIHSAKHKVDICQKRVVVYDOSQDVASLSDCFLLVLLGLKESFNSVHL 120
 DB 61 DKVLITELIHSAKHKVDICQKRVVVYDOSQDVASLSDCFLLVLLGLKESFNSVHL 120

QY 121 LAGFAEFSRCFPGLCCEGKSTLVPTCISQPCLPVANIGPRLIPNLVLCGQDVLNKL 180
 DB 121 LAGFAEFSRCFPGLCCEGKSTLVPTCISQPCLPVANIGPRLIPNLVLCGQDVLNKL 180
 QY 181 QQNGIGVLMASVTCRPFDPFIPESHFLRPVNDSCFEKILPWLKSVDFTEKASNGCV 240
 DB 181 QQNGIGVLMASVTCRPFDPFIPESHFLRPVNDSCFEKILPWLKSVDFTEKASNGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKRDMSIDENYRVKERRPTISNPNFLGOLLDYKKIKN 300
 DB 241 LVHCLAGISRSATIAIAYIMKRDMSIDENYRVKERRPTISNPNFLGOLLDYKKIKN 300
 QY 301 QTAGASGPKSKLKLHLKPNRPVAVSEGGQKSETPSPCCASATSSEAAGORPVHPASV 360
 DB 301 QTAGASGPKSKLKLHLKPNRPVAVSEGGQKSETPSPCCASATSSEAAGORPVHPASV 360
 QY 361 PSVPSVQPSLLEDSPVLVALSGHLASDRLEDENKLRSPSLDIKSVYSASMAASLHGF 420
 DB 361 PSVPSVQPSLLEDSPVLVALSGHLASDRLEDENKLRSPSLDIKSVYSASMAASLHGF 420
 QY 421 SSSSEDALEYKYSTTIDGTNKLCOFSPVQELASQTEPSTPDKEASIPKQLQTPAPSDQ 480
 DB 421 SSSSEDALEYKYSTTIDGTNKLCOFSPVQELASQTEPSTPDKEASIPKQLQTPAPSDQ 480
 QY 481 SKRLHSVRTSSSGTAORSLSPLRHSGSVEDNHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
 DB 481 SKRLHSVRTSSSGTAORSLSPLRHSGSVEDNHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
 QY 541 SDILAPQTSPLTSSSWFYATBESSHFFYSASAIYGSASVAYSCSQLPTCGDQVYVRRR 600
 DB 541 SDILAPQTSPLTSSSWFYATBESSHFFYSASAIYGSASVAYSCSQLPTCGDQVYVRRR 600
 QY 601 QKPSDPAISRHSWHESPPFKQKRRSCOMEFESITSENRREBELGKVGSSQSFSGSME 660
 DB 601 QKPSDPAISRHSWHESPPFKQKRRSCOMEFESITSENRREBELGKVGSSQSFSGSME 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665
 RESULT 6
 AAU09946
 ID AAU09946 standard; protein; 665 AA.
 AC AAU09946;
 DT 18-JUN-2002 (first entry)
 DE Protein sequence of human (dual specificity phosphatase) DSP-10.
 XX
 XX Human; dual specificity phosphatase; DSP-10; cancer; epilepsy; stroke;
 KW neuronal degeneration syndrome; Alzheimer's disease; depression;
 KW schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
 KW osteoporosis; diabetes.
 XX
 XX Homo sapiens.
 OS
 XX
 PN MO200177340-A1.
 XX
 PD 18-OCT-2001.
 PF 06-APR-2001; 2001MO-EP003966.
 XX
 PR 10-APR-2000; 2000EP-00107143.
 XX
 PA (MERE) MERCK PATENT GMBH.
 PI Duecker K;
 XX
 XX WPI; 2002-010917/01.
 DR N-PSDB; AAS15768.
 XX

PT Novel dual specificity phosphatase polypeptides useful for treating
PT cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
PT disease, depression, schizophrenia, asthma and immune disorders.
XX
PS Claim 2; Page 37-39; 43pp; English.

CC The present invention relates to a new isolated dual specificity
CC phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
CC sequence that is fully defined in the specification. The invention also
CC provides a sequence encoded by a 3059 nucleotide sequence fully defined
CC in the specification, and a sequence having at least 95 % identity to the
CC polypeptide, or fragments or variants of DUSP-10. The invention is useful
CC for treating cancer e.g. leukaemia, colon carcinoma, lung cancer,
CC prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
CC stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
CC schizophrenia, cardiac myopathies, asthma, immune disorders,
CC inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
CC osteoporosis, diabetes and diabetes associated diseases. The molecules of
CC the invention are also useful as vaccines for inducing immunological
CC response in a mammal, in disease diagnosis and in assays for screening
CC agonistic or antagonistic compounds. Other uses of the invention include
CC identifying membrane bound or soluble receptors, as a diagnostic reagent,
CC in chromosome localization studies, and as a valuable tool in tissue
CC expression studies. The present sequence represents the dual specificity
CC phosphatase, DUSP-10, protein of the invention

XX Sequence 665 AA;

Query Match 99.6%; Score 3406; DB 5; Length 665;
Best Local Similarity 99.7%; Pred. No. 2,76-256;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLSEGTETKVLIDSRPVENYTHILLEANINCSKLMKRLQQ 60
DB 1 MAHEMIGTQVTERLVALLSEGTETKVLIDSRPVENYTHILLEANINCSKLMKRLQQ 60
QY 61 DKVLITELIHSAGKAVDIDCSOKRVVYDOSDVASLSDDCLVTLGLKLEKSNVHL 120
DB 61 DKVLITELIHSAGKAVDIDCSOKRVVYDOSDVASLSDDCLVTLGLKLEKSNVHL 120
QY 121 LAGGFAEFRCFPELCEGKSTLVPTCISQPCLPVANTIGTRLLPMLYIGCCQDVLNKLKI 180
DB 121 LAGGFAEFRCFPELCEGKSTLVPTCISQPCLPVANTIGTRLLPMLYIGCCQDVLNKLKI 180
QY 121 LAGGFAEFRCFPELCEGKSTLVPTCISQPCLPVANTIGTRLLPMLYIGCCQDVLNKLKI 180
DB 121 LAGGFAEFRCFPELCEGKSTLVPTCISQPCLPVANTIGTRLLPMLYIGCCQDVLNKLKI 180
QY 181 QONGIGVYVMASTTCRCPDPIFESHFLRVVNDSCFKILPMLDKSVDFIKAKASNGCV 240
DB 181 QONGIGVYVMASTTCRCPDPIFESHFLRVVNDSCFKILPMLDKSVDFIKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIYIMKMDMSLBEAYRFVYKERTTISPNNFLGQILLDYEKIKN 300
DB 241 LVHCLAGISRSATIAIYIMKMDMSLBEAYRFVYKERTTISPNNFLGQILLDYEKIKN 300
QY 301 QTCASGPRKQLHLHKEKNEPVPVAVSBGQKSETPLSPPCADSATSEAGORPVHPASV 360
DB 301 QTCASGPRKQLHLHKEKNEPVPVAVSBGQKSETPLSPPCADSATSEAGORPVHPASV 360
QY 361 PSVPVQPSLLEDSPLVOALSGHLHSDRLSDSKLKRSTSLDIKSVYSASMAASLHGF 420
DB 361 PSVPVQPSLLEDSPLVOALSGHLHSDRLSDSKLKRSTSLDIKSVYSASMAASLHGF 420
QY 421 SSSBNALEYRPTTLTGNTKLCOFSPVOELSEQTPETSPDKXEAIPKKLQTPARSDSQ 480
DB 421 SSSBNALEYRPTTLTGNTKLCOFSPVOELSEQTPETSPDKXEAIPKKLQTPARSDSQ 480
QY 481 SKRLHVSRTSSGTAQSRSLSPHRSQSVNDNHTSFLGLSTSQOHLTKSAGLGLKGMH 540
DB 481 SKRLHVSRTSSGTAQSRSLSPHRSQSVNDNHTSFLGLSTSQOHLTKSAGLGLKGMH 540
QY 541 SDLIAPOTPSLTSSMYFATSSHFYSAIYGSASYSAYSCGOLPTCGOVYSVRR 600
DB 541 SDLIAPOTPSLTSSMYFATSSHFYSAIYGSASYSAYSCGOLPTCGOVYSVRR 600
QY 601 QKPSDRADSRHSRHSPEFKQFKRRSCOMERGESIMENSRBELGKVGSSPFGSME 660
DB 601 QKPSDRADSRHSRHSPEFKQFKRRSCOMERGESIMENSRBELGKVGSSPFGSME 660

DB 601 QKPSDRADSRHSRHSPEFKQFKRRSCOMERGESIMENSRBELGKVGSSPFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665
RESULT 7
AAU75789
ID AAU75789 standard; protein; 665 AA.
AC AAU75789;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human protein phosphatase 7 (PP7) protein sequence.
XX
KW Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;
KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
KW melanoma; myeloma sarcoma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Domain 11..131
FT /label=Rhodanese_like_domain
FT 15..170
FT /note="Dual specificity protein phosphatase"
FT 85..298
FT /note="VH1-type dual specificity protein phosphatase"
FT 158..297
FT /label=Catalytic domain_DSPC
FT /note="Dual specificity phosphatase"
FT 220..281
FT Region
FT site="Tyrosine specific protein phosphatases active
FT /note="Tyrosine specific protein phosphatases active
FT 242..254
FT /note="Tyrosine phosphatase"
FT 242..252
FT /note="Tyrosine specific protein phosphatases"
XX
PD WO200210363-A2.
XX
PD 07-FEB-2002.
XX
PF 26-JUL-2001; 2001WO-US023716.
XX
PR 28-JUL-2000; 2000US-0221679P.
PR 03-AUG-2000; 2000US-0223372P.
PR 10-AUG-2000; 2000US-0224309P.
PR 18-AUG-2000; 2000US-0226728P.
PR 30-AUG-2000; 2000US-0229254P.
PR 08-SEP-2000; 2000US-0231366P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Tang YT, Elliott VS, Rankumar J, Yao MG, Burford N, Wang YE;
PI Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalia AJA, Lu DAM;
PI Tribouley CM, Griffin JA, Baughn MK, Yue H, Warren BA, Nguyen DB;
PI Malia NK, Kearney L;
DR WPI; 2002-188735/24.
DR N-PSDB; ABK14474.
XX
PT New protein phosphatases, useful for diagnosing, treating or preventing
PT immune system disorders (e.g. Crohn's disease), neurological disorders
PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
PT cancers).
XX

PS Claim 1; Page 107-108; 117pp; English.

XX The present invention relates to a new polypeptide, a naturally occurring
CC amino acid sequence at least 95 % identical to it, a biologically active
CC fragment of it or an immunogenic fragment of it. The polypeptides,
CC polynucleotides, agonists and antagonists are useful for diagnosing,
CC treating or preventing disorders associated with aberrant expression of
CC protein phosphatases (PP), particularly immune system disorders e.g.
CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
CC achmia or Crohn's disease, neurological disorders e.g. epilepsy,
CC Huntington's disease, dementia or Parkinson's disease, developmental
CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
CC or sarcoma. The present amino acid sequence represents human protein
CC phosphatase 7 (PP7) which is one of several human protein phosphatases
CC (AAU5783-AAU5792) of the invention

XX Sequence 665 AA;

Query Match 99.6%; Score 3406; DB 5; Length 665;

Best Local Similarity 99.7%; Pred. No. 2.7e-256;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRRLQ 60
QY 61 DKVLITELLQHSAAKHVDIDCSQKVVYDQSSQDVASLSSDCFLTVLLGLKESFNSVHL 120
DB 61 DKVLITELLQHSAAKHVDIDCSQKVVYDQSSQDVASLSSDCFLTVLLGLKESFNSVHL 120
QY 121 LAGGFAPSPRCFPGLCGKSTLVPTCISQPCLPVANIPTRIIPNLVYLGQRDVNLKELI 180
DB 121 LAGGFAPSPRCFPGLCGKSTLVPTCISQPCLPVANIPTRIIPNLVYLGQRDVNLKELI 180
QY 181 QONGIGYVLANASTCPKPDFIPESHFLRVVNDSPCEKILPMWLDKSDVFIEKKAANGCV 240
DB 181 QONGIGYVLANASTCPKPDFIPESHFLRVVNDSPCEKILPMWLDKSDVFIEKKAANGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRTISPNNFNLQGLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRTISPNNFNLQGLLDYEKKIKN 300
QY 301 QTASGPKSKLKLHLKEKNEPVPVASEGGQSETPSPCADSATSBAAGRPVHPASV 360
DB 301 QTASGPKSKLKLHLKEKNEPVPVASEGGQSETPSPCADSATSBAAGRPVHPASV 360
QY 361 PSVPSPVPSLLEDSPLVQALSGHLASDRLEDSNKLKSPSLDIKSVYSASMAASLHGF 420
DB 361 PSVPSPVPSLLEDSPLVQALSGHLASDRLEDSNKLKSPSLDIKSVYSASMAASLHGF 420
QY 421 SSSSEDALEYKPSSTLLDGTNKLQPSFVQBLSEQTPEISPDKEEASIPKKLQIARPSDQ 480
DB 421 SSSSEDALEYKPSSTLLDGTNKLQPSFVQBLSEQTPEISPDKEEASIPKKLQIARPSDQ 480
QY 481 SKRLHSVRRSSSGTAQRSLSPHRSQSVEDNYHTSFLFGLSTSQOHLTSAGLGLKGM 540
DB 481 SKRLHSVRRSSSGTAQRSLSPHRSQSVEDNYHTSFLFGLSTSQOHLTSAGLGLKGM 540
QY 541 SDLLAQTSTPSLTSSMYFATESSHFASAGAIYGSSAYASCSOLPTGDDVYSVRR 600
DB 541 SDLLAQTSTPSLTSSMYFATESSHFASAGAIYGSSAYASCSOLPTGDDVYSVRR 600
QY 601 QKPSDRAISRMSHESPEFKQFKRRSCQMEFGESIMSENRREBELGKVSGSSFGSME 660
DB 601 QKPSDRAISRMSHESPEFKQFKRRSCQMEFGESIMSENRREBELGKVSGSSFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 8

ABB97946
ID ABB97946 standard; protein; 665 AA.

XX ABB97946;

XX 06-SEP-2002 (first entry)

XX Human protein sequence #13.

XX Human; brain; tonsil; hippocampus; foetal brain; diagnosis.

XX Homo sapiens.

XX W0200252005-A1.

XX 04-JUL-2002.

XX 20-DEC-2001; 2001WO-JP011217.

XX 22-DEC-2000; 2000JP-00389742.

XX (KAZU-) KAZUSA DNA RES INST FOUND.

XX (CELE-) CELESTAR LEXICO-SCI LTD.

XX Ohara O, Nagase T, Nakajima D;

XX WPI, 2002-500762/53.

XX N-PSDB; ABB93966.

XX Genes and their expression products cloned from human cDNA libraries for
XX treatment and diagnosis of diseases associated with their expression.

XX Claim 1(a); Page 112-116; 238pp; Japanese.

XX The invention relates to DNA encoding polypeptides directly cloned from
XX cDNA libraries originating in adult whole brain, human tonsil, human
XX adult hippocampus and human foetal brain. Polypeptides and
XX polynucleotides of the invention may be used in the investigation of
XX differential expression of the DNA sequences in normal subjects and
XX disease patients. They may also be used in the production of antibodies,
XX oligonucleotide probes and DNA chips for diagnosis and identification of
XX drugs for treatment of diseases with which the DNA sequences are
XX associated. The sequences given in records ABB97934-A-BB97964 represent
XX human proteins of the invention

XX Sequence 665 AA;

Query Match 99.6%; Score 3406; DB 5; Length 665;

Best Local Similarity 99.7%; Pred. No. 2.7e-256;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRRLQ 60
QY 61 DKVLITELLQHSAAKHVDIDCSQKVVYDQSSQDVASLSSDCFLTVLLGLKESFNSVHL 120
DB 61 DKVLITELLQHSAAKHVDIDCSQKVVYDQSSQDVASLSSDCFLTVLLGLKESFNSVHL 120
QY 121 LAGGFAPSPRCFPGLCGKSTLVPTCISQPCLPVANIPTRIIPNLVYLGQRDVNLKELI 180
DB 121 LAGGFAPSPRCFPGLCGKSTLVPTCISQPCLPVANIPTRIIPNLVYLGQRDVNLKELI 180
QY 181 QONGIGYVLANASTCPKPDFIPESHFLRVVNDSPCEKILPMWLDKSDVFIEKKAANGCV 240
DB 181 QONGIGYVLANASTCPKPDFIPESHFLRVVNDSPCEKILPMWLDKSDVFIEKKAANGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRTISPNNFNLQGLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRTISPNNFNLQGLLDYEKKIKN 300
QY 301 QTASGPKSKLKLHLKEKNEPVPVASEGGQSETPSPCADSATSBAAGRPVHPASV 360
DB 301 QTASGPKSKLKLHLKEKNEPVPVASEGGQSETPSPCADSATSBAAGRPVHPASV 360

```

Db      301 QTGASGPKSKLHLLEKNEPVPVAVSEGGOKSTPLSPCCADSAISEAAGRPVPAV
Qy      361 PSVPVQPSLLEDPPLVQALSGHLSDRLSDNSKLRSPSLDIKSYSASMAASLHG
Db      361 PSVPVQPSLLEDPPLVQALSGHLSDRLSDNSKLRSPSLDIKSYSASMAASLHG
Qy      421 SSSDALERYKPSLTLDGTNKLCOFSPVQELSDQTEGSPDKEASIPKQLQTA
Db      421 SSSDALERYKPSLTLDGTNKLCOFSPVQELSDQTEGSPDKEASIPKQLQTA
Qy      481 SKRLHVRSSSGTAQOQSLSLPLHRSQVYEDNHTSFLPGLSTSOQHLTKSAGLKG
Db      481 SKRLHVRSSSGTAQOQSLSLPLHRSQVYEDNHTSFLPGLSTSOQHLTKSAGLKG
Qy      541 SDILAPQTSPLTSSWYFATESSHFYSAIYGSASASNSCQLPTCGDYVSVRR
Db      541 SDILAPQTSPLTSSWYFATESSHFYSAIYGSASASNSCQLPTCGDYVSVRR
Qy      601 QKPSDADRRSRHESHPPEKOPKRRSCMEFGEISMEENSRRELGKVGSSFSG
Db      601 QKPSDADRRSRHESHPPEKOPKRRSCMEFGEISMEENSRRELGKVGSSFSG
Qy      661 IIEVS 665
Db      661 IIEVS 665

```

```

RESULT 9
AAU79929
AAU79929 standard; protein; 665 AA.

```

```

AC      AAU79929;
XX      02-JUL-2002 (first entry)
DE      Human dual specificity phosphatase 21117 protein.
XX      Human dual specificity phosphatase 21117; erythroid-related disorder;
KW      haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
KV      erythrocytosis; liver-related disorder; cancer.
XX      Homo sapiens.
OS
XX
FH      Key
FH      Domain
FT      /label= Rhodanese_like_domain
FT      21..24
FT      /note= "Casein kinase II phosphorylation site"
FT      91..94
FT      /note= "Casein kinase II phosphorylation site"
FT      158..297
FT      /note= "Dual specificity phosphatase catalytic domain"
FT      214..217
FT      /note= "Casein kinase II phosphorylation site"
FT      242..254
FT      /note= "Tyrosine specific protein phosphatase active
FT      site"
FT      242..254
FT      /note= "C-X5-R motif"
FT      266..269
FT      /note= "Casein kinase II phosphorylation site"
FT      369..372
FT      /note= "Casein kinase II phosphorylation site"
FT      421..424
FT      /note= "Casein kinase II phosphorylation site"
FT      434..437
FT      /note= "Casein kinase II phosphorylation site"
FT      458..461
FT      /note= "Casein kinase II phosphorylation site"
FT      508..511
FT      /note= "Casein kinase II phosphorylation site"
FT      589..592
FT      /note= "Casein kinase II phosphorylation site"
FT      /note= "Casein kinase II phosphorylation site"

```

```

FT      Region
FT      /note= "Casein kinase II phosphorylation site"
FT      612..615
FT      Region
FT      /note= "Casein kinase II phosphorylation site"
FT      617..620
FT      Region
FT      /note= "Casein kinase II phosphorylation site"
FT      642..645
FT      /note= "Casein kinase II phosphorylation site"
PN      US2002034807-A1.
XX      21-MAR-2002.
XX      23-MAR-2001; 2001US-00816494.
XX      24-MAR-2000; 2000US-01918589.
XX      (MEYER/ MEYERS R A.
XX      Meyers RA;
XX      WPI; 2002-351086/38.
XX      N-PSDB; ABK49402.
XX
XX      New nucleic acids, designated 38692 and 21117, encoding dual specificity
XX      phosphatases for treating cell proliferation and differentiation
XX      disorders including hematopoietic and erythroid-related disorders and
XX      cancers.
XX      Claim 8; Fig 1; 76pp; English.
XX
XX      The present invention relates to new nucleic acids designated 38692 and
XX      21117 encoding dual specificity phosphatase family members. The nucleic
XX      acid, polypeptide encoded by it, and antibody specific for the
XX      polypeptide may be used to diagnose and treat haematopoietic-related
XX      disorders such as leukaemias and autoimmune diseases, erythroid-related
XX      disorders and cancers, particularly of the breast, colon, adipose, prostate and
XX      lung. The present amino acid sequence represents the human dual
XX      specificity phosphatase 21117 protein of the invention, as described
XX      above
XX
SQ      Sequence 665 AA;

```

```

Qy      1 MAHEMIGQVYTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMRRLOQ
Db      1 MAHEMIGQVYTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMRRLOQ
Qy      61 DKVLITELIOHSAKHKVDIDCGKVVVYDSSODVASLSDCFVLVLLGKLEKSFNSVHL
Db      61 DKVLITELIOHSAKHKVDIDCGKVVVYDSSODVASLSDCFVLVLLGKLEKSFNSVHL
Qy      121 LAGGFAEFSCFPGLOEGKSTLVPTCISQCLPVANIGPTILPVLVYGCQDVYANKELI
Db      121 LAGGFAEFSCFPGLOEGKSTLVPTCISQCLPVANIGPTILPVLVYGCQDVYANKELI
Qy      181 QONGIGYVLANASYTCRPFIPESHFLRVVNDSCETILPMWLDKSVDFIEKASNGCV
Db      181 QONGIGYVLANASYTCRPFIPESHFLRVVNDSCETILPMWLDKSVDFIEKASNGCV
Qy      241 LVHCLAGISRSATVIAIYIMKRMDSIDEAYRPFYKERTIPSPNFNLGQLDYEEKIKN
Db      241 LVHCLAGISRSATVIAIYIMKRMDSIDEAYRPFYKERTIPSPNFNLGQLDYEEKIKN
Qy      301 QTGASGPKSKLHLLEKNEPVPVAVSEGGOKSTPLSPCCADSAISEAAGRPVPAV
Db      301 QTGASGPKSKLHLLEKNEPVPVAVSEGGOKSTPLSPCCADSAISEAAGRPVPAV
Qy      361 PSVPVQPSLLEDPPLVQALSGHLSDRLSDNSKLRSPSLDIKSYSASMAASLHG
Db      361 PSVPVQPSLLEDPPLVQALSGHLSDRLSDNSKLRSPSLDIKSYSASMAASLHG

```



```

QY 421 SSEDALERYKSTTLDGNTKLCQFSPVOELSEOTPETSPDKKEASIPKKLQTPARPSDSQ 480
Db 421 SSEDALERYKSTTLDGNTKLCQFSPVOELSEOTPETSPDKKEASIPKKLQTPARPSDSQ 480
QY 481 SKRLHSVRTSSSGTAORSLSLPLHRSGSVEDNYHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
Db 481 SKRLHSVRTSSSGTAORSLSLPLHRSGSVEDNYHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
QY 541 SDILAPQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDQVYSVRRR 600
Db 541 SDILAPQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDQVYSVRRR 600
QY 601 QKPSDADSRSSWHESSPFEKQFKRRSCOMFEGESIMSENRSRBELGKVGSSQSSFGSME 660
Db 601 QKPSDADSRSSWHESSPFEKQFKRRSCOMFEGESIMSENRSRBELGKVGSSQSSFGSME 660
QY 661 IIEVS 665
Db 661 IIEVS 665

```

RESULT 10

ABB97291
ID ABB97291 strand; protein; 665 AA.

AC ABB97291;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 559.

KM Human; anti-naemic; vulnerary; anti-inflammatory; immunomodulator;
KW anti-infective; cerebroprotective; cyostatic; rheumatic; gene therapy;
KM neuroprotective; anti-parkinsonian; protein therapy; EST;
XX expressed sequence tag.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001MO-US026015.

PR 11-SEP-2000; 2000US-00659671.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Maundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Dirmann RT;

DR WPI; 2002-292408/33.

DR N-PSDB; ABN32477.

PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.

PS Example 2; SEQ ID NO 559; 509bp; English.

PS The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibin e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke

CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.

CC rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention

Query Match 99.6%; Score 3406; DB 5; Length 665;
Best Local Similarity 99.7%; Pred. No. 2,7e-256;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MAHEMIGTOIVTERVALLSEGTAKVLLIDSRPFVYNTSHIEAININCSKIMKRLQ 60
Db 1 MAHEMIGTOIVTERVALLSEGTAKVLLIDSRPFVYNTSHIEAININCSKIMKRLQ 60
QY 61 DKVLITELLQHSKAKHVDIDCQKVVVYVDSQDVASLSDDCLTVLLGLKERSFMSVHL 120
Db 61 DKVLITELLQHSKAKHVDIDCQKVVVYVDSQDVASLSDDCLTVLLGLKERSFMSVHL 120
QY 121 LAGGFARFCRPGCLCEGSTLVPTCISQPCLPVANIIGTRILPNIYLGCGQVRLKELM 180
Db 121 LAGGFARFCRPGCLCEGSTLVPTCISQPCLPVANIIGTRILPNIYLGCGQVRLKELM 180
QY 181 QQNGIGYVLANASTYCPKDPFIPESHELRVPVNDSPCEKILPMWIDKSVDFIEKAKANGCV 240
Db 181 QQNGIGYVLANASTYCPKDPFIPESHELRVPVNDSPCEKILPMWIDKSVDFIEKAKANGCV 240
QY 241 LVHCLAGISRATIAIYIMKMDNSLDRAVRFVKEKRTISPENFQLQLDYEKKIKN 300
Db 241 LVHCLAGISRATIAIYIMKMDNSLDRAVRFVKEKRTISPENFQLQLDYEKKIKN 300
QY 301 QTGASGPKSKLKLHLEKNEBPVAVSEGGOKSETPLSPCADSATSEAAQRPVHPASV 360
Db 301 QTGASGPKSKLKLHLEKNEBPVAVSEGGOKSETPLSPCADSATSEAAQRPVHPASV 360
QY 361 PSVPSPVQPSLLEDSPLVQALSGHLASADRLLEDNSNKLKRSFLDIKSVYSASMAASLHGF 420
Db 361 PSVPSPVQPSLLEDSPLVQALSGHLASADRLLEDNSNKLKRSFLDIKSVYSASMAASLHGF 420
QY 421 SSEDALERYKSTTLDGNTKLCQFSPVOELSEOTPETSPDKKEASIPKKLQTPARPSDSQ 480
Db 421 SSEDALERYKSTTLDGNTKLCQFSPVOELSEOTPETSPDKKEASIPKKLQTPARPSDSQ 480
QY 481 SKRLHSVRTSSSGTAORSLSLPLHRSGSVEDNYHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
Db 481 SKRLHSVRTSSSGTAORSLSLPLHRSGSVEDNYHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
QY 541 SDILAPQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDQVYSVRRR 600
Db 541 SDILAPQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDQVYSVRRR 600
QY 601 QKPSDADSRSSWHESSPFEKQFKRRSCOMFEGESIMSENRSRBELGKVGSSQSSFGSME 660
Db 601 QKPSDADSRSSWHESSPFEKQFKRRSCOMFEGESIMSENRSRBELGKVGSSQSSFGSME 660
QY 661 IIEVS 665
Db 661 IIEVS 665

```

RESULT 11

ID ADA54744 strand; protein; 665 AA.

AC ADA54744;

DT 20-NOV-2003 (first entry)

DE Human protein, SEQ ID 2312.

KM Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KM inflammatory disease; osteoporosis; neurological disease.

OS Homo sapiens.

PN BP1293569-A2.

PD 19-MAR-2003.

PF 21-MAR-2002; 2002EP-00006586.
 XX
 PR 14-SEP-2001; 2001JP-00328381.
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Irie R, Tamechika I,
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
 DR WPI; 2003-395539/38.
 DR N-PSDB; ADA53105.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 XX
 PS Claim 14; SEQ ID NO 2312; 205pp; English.

Sequence 665 AA;

Query Match 99.4%; Score 3399; DB 6; Length 665;
 Best Local Similarity 99.5%; Pred. No. 9,4e-256;
 Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTEVRLVALLSGTEKYLIDSRPEVYNTSHILEAININCSKLMRRRLQ 60
 DB 1 MAHEMIGTQVTEVRLVALLSGTEKYLIDSRPEVYNTSHILEAININCSKLMRRRLQ 60
 QY 61 DKVLTIELIQHSAKHKVIDCSQKVVYVDSQDVASLSDCEFLVTLGKLEKSNVYL 120
 DB 61 DKVLTIELIQHSAKHKVIDCSQKVVYVDSQDVASLSDCEFLVTLGKLEKSNVYL 120
 QY 121 LAGGFAEFSRCFPGLECGKSTVPTCTISOPCLPVANIGFTRKLPVLYGCGQDVYLNKELI 180
 DB 121 LAGGFAEFSRCFPGLECGKSTVPTCTISOPCLPVANIGFTRKLPVLYGCGQDVYLNKELI 180
 QY 121 LAGGFAEFSRCFPGLECGKSTVPTCTISOPCLPVANIGFTRKLPVLYGCGQDVYLNKELI 180
 DB 121 LAGGFAEFSRCFPGLECGKSTVPTCTISOPCLPVANIGFTRKLPVLYGCGQDVYLNKELI 180
 QY 181 QONGIGYVLAATYTCCKPDPIFESHFLRVVDSFCCKILPMLDKSVDFIEKASNGCV 240
 DB 181 QONGIGYVLAATYTCCKPDPIFESHFLRVVDSFCCKILPMLDKSVDFIEKASNGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKRPPTISPNFNLGQLLDYEKKIKN 300
 DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKRPPTISPNFNLGQLLDYEKKIKN 300
 QY 301 QGASGPGSKLTLLEKNEPVPVAVSGQKSETPSPPCADATSEAGORPVHNAV 360
 DB 301 QGASGPGSKLTLLEKNEPVPVAVSGQKSETPSPPCADATSEAGORPVHNAV 360
 QY 361 PSVPVQSLLEDSPLVQALSGHLISADRLSDSKLKRSTSLDIKSVYSASVAASLHGF 420
 DB 361 PSVPVQSLLEDSPLVQALSGHLISADRLSDSKLKRSTSLDIKSVYSASVAASLHGF 420
 QY 421 SSSSDALRYRPTSTLLDGTNKLCOFSPVQELSEOTPEPSPDKREASIPKKLQTARPSDQ 480
 DB 421 SSSSDALRYRPTSTLLDGTNKLCOFSPVQELSEOTPEPSPDKREASIPKKLQTARPSDQ 480
 QY 481 SKRLHVRITSSSGAQRSLSLPLHRSVYNDNTHTSFLGLSTQOHLTKSAGLAKGMH 540
 DB 481 SKRLHVRITSSSGAQRSLSLPLHRSVYNDNTHTSFLGLSTQOHLTKSAGLAKGMH 540
 QY 541 SDLIAPGTSTPSLTSSWYPATESHRYFASAIYGSASVAYSCGOLPTCGQVYVARR 600
 DB 541 SDLIAPGTSTPSLTSSWYPATESHRYFASAIYGSASVAYSCGOLPTCGQVYVARR 600

QY 601 QKPSDRAISRGRHSESPFEKQFKRRSCOMERGESIMSENRREELKGVGSSQSFSGSME 660
 DB 601 QKPSDRAISRGRHSESPFEKQFKRRSCOMERGESIMSENRREELKGVGSSQSFSGSME 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 12

AAU79161 standard; protein; 665 AA.
 AAU79161;

02-JUL-2002 (first entry)

Human dual-specificity phosphatase-3 (DSP-16) mutant protein #1.

Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 cancer; graft-versus-host disease; allergy; metabolic disease;
 abnormal cell growth; abnormal cell proliferation; contact inhibition;
 cell cycle abnormality; anchorage independent cell growth; apoptosis;
 intercellular adhesion; DSP-16 modulator; mutant; mutein.

Homo sapiens.
 Synthetic.

Key Location/Qualifiers
 Misc-difference 213 /note= "Wild-type Asp substituted by Ala"

W0200226997-A2.

04-APR-2002.

25-SEP-2001; 2001WO-US030124.

26-SEP-2000; 2000US-0235487P.

(CEPT-) CEPTYR INC.

Luche RM, Wei B;

WPI; 2002-315802/35.

New DSP-16 polypeptide, useful for identifying modulators of its
 activity, which can be used in the treatment of disorders such as
 Duchenne muscular dystrophy, or cancer.

Claim 46; Page; 87pp; English.

The present invention relates to a new polypeptide, DSP-16, having a 665
 amino acid sequence, given in the specification, or a variant having at
 least 50 % identical residues, which retains the ability to
 dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 invention can be used for identifying agents which modulate DSP-16
 activity, for modulation of a proliferative response in a cell, survival
 of a cell, or differentiation of a cell. The cell displays contact
 inhibition of cell growth or anchorage independent growth and may display
 altered intercellular adhesion. The agent may modulate apoptosis, or the
 cell cycle. The identified modulators can be used to treat Duchenne
 muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 cell proliferation, and cell cycle abnormalities. The present amino acid
 sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 mutant protein #1. Note: This sequence is not shown in the specification
 in figure 2 of the specification

Sequence 665 AA;

Query Match 99.4%; Score 3398; DB 5; Length 665;
 Best Local Similarity 99.5%; Pred. No. 1,1e-255;
 Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMIGTOIVTERLVALLESGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
 DB 1 MAHEMIGTOIVTERLVALLESGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
 QY 61 DKVLITELIOWSAKHKVDIDCSQKVVVYDSSQDVASLSDCFLTVLLGLKLEKSFNSVHL 120
 DB 61 DKVLITELIOWSAKHKVDIDCSQKVVVYDSSQDVASLSDCFLTVLLGLKLEKSFNSVHL 120
 QY 121 LAGGFAEFRCFPGLCGKSTLVPTCISQPCLPVANIIGPRRIIPNLYLGGQDVLMKELI 180
 DB 121 LAGGFAEFRCFPGLCGKSTLVPTCISQPCLPVANIIGPRRIIPNLYLGGQDVLMKELI 180
 QY 181 QONGIGVYLNASTCPKPDPIPSHFLRVVNDSPCEKILPMDKSVDFLEKAKASNGCV 240
 DB 181 QONGIGVYLNASTCPKPDPIPSHFLRVVNDSPCEKILPMDKSVDFLEKAKASNGCV 240
 QY 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKN 300
 DB 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKN 300
 QY 301 QTGASGPKSKLKLHLEKNEPVAVSEGGQKSTPLSPCADSATSEAGORPVHPASV 360
 DB 301 QTGASGPKSKLKLHLEKNEPVAVSEGGQKSTPLSPCADSATSEAGORPVHPASV 360
 QY 361 PSVPVQPSLLEDSPLVQALSGHLASADRLSDNSKLRSPSLDIKSVYSASMAASLHGF 420
 DB 361 PSVPVQPSLLEDSPLVQALSGHLASADRLSDNSKLRSPSLDIKSVYSASMAASLHGF 420
 QY 421 SSSBDALYYKPTTLDGTNKLCPSPVQELSEQTPETSVDKEBASIPKLLQTPARPSDQ 480
 DB 421 SSSBDALYYKPTTLDGTNKLCPSPVQELSEQTPETSVDKEBASIPKLLQTPARPSDQ 480
 QY 481 SKRLHSVRTSSGAGQKSLSPHRSQSVEDNYHTSLFGLSTSQQLTASAGLKGWH 540
 DB 481 SKRLHSVRTSSGAGQKSLSPHRSQSVEDNYHTSLFGLSTSQQLTASAGLKGWH 540
 QY 541 SDLIAPOTSPSLTSSWYFATESHFFSASAIYCGSAYAVSCQPLTCGDQVYVRRR 600
 DB 541 SDLIAPOTSPSLTSSWYFATESHFFSASAIYCGSAYAVSCQPLTCGDQVYVRRR 600
 QY 601 QKPSDRADSRSSWHESSPEKQKRRSCQMEFGESIMENSRRELKGVSGSSFGSME 660
 DB 601 QKPSDRADSRSSWHESSPEKQKRRSCQMEFGESIMENSRRELKGVSGSSFGSME 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 13
 AAU79162 standard; protein; 665 AA.
 AAU79162;
 02-JUL-2002 (first entry)

DE Human dual-specificity phosphatase-3 (DSP-16) mutant protein #2.
 KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; mutant; mutcin.
 XX Homo sapiens.
 OS Synthetic.
 XX

FH Key Location/Qualifiers
 FT Misc-difference 244 /note="Wild-type Cys substituted by Ser"
 PN MO200226997-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 25-SEP-2001; 2001MO-US030124.
 XX
 PR 26-SEP-2000; 2000US-0235487P.
 XX
 PA (CEPT-) CEPTYR INC.
 XX
 PI Luche RM, Wei B;
 XX
 PS WPI; 2002-315802/35.
 DR
 XX
 PT New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX
 XX Claim 46; Page; 87pp; English.

The present invention relates to a new polypeptide, DSP-16, having a 665 amino acid sequence, given in the specification, or a variant having at least 50 % identical residues, which retains the ability to dephosphorylate an activated mitogen-activated protein (MAP) kinase. The invention can be used for identifying agents which modulate DSP-16 activity, for modulation of a proliferative response in a cell, survival of a cell, or differentiation of a cell. The cell displays contact inhibition of cell growth or anchorage independent growth and may display altered intercellular adhesion. The agent may modulate apoptosis, or the cell cycle. The identified modulators can be used to treat Duchenne muscular dystrophy, cancer, graft-versus-host disease, autoimmune diseases, allergies, metabolic diseases, abnormal cell growth, abnormal cell proliferation, and cell cycle abnormalities. The present amino acid sequence represents the human dual-specificity phosphatase-3 (DSP-16) mutant protein #2. Note: This sequence is not shown in the specification but is derived from the wild-type human DSP-16 (AAU79156) protein given in figure 2 of the specification

Sequence 665 AA:

Query Match 99.4%; Score 3396; DB 5; Length 665;
 Best Local Similarity 99.5%; Pred. No. 1,1e-255;
 Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMIGTOIVTERLVALLESGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
 DB 1 MAHEMIGTOIVTERLVALLESGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
 QY 61 DKVLITELIOWSAKHKVDIDCSQKVVVYDSSQDVASLSDCFLTVLLGLKLEKSFNSVHL 120
 DB 61 DKVLITELIOWSAKHKVDIDCSQKVVVYDSSQDVASLSDCFLTVLLGLKLEKSFNSVHL 120
 QY 121 LAGGFAEFRCFPGLCGKSTLVPTCISQPCLPVANIIGPRRIIPNLYLGGQDVLMKELI 180
 DB 121 LAGGFAEFRCFPGLCGKSTLVPTCISQPCLPVANIIGPRRIIPNLYLGGQDVLMKELI 180
 QY 181 QONGIGVYLNASTCPKPDPIPSHFLRVVNDSPCEKILPMDKSVDFLEKAKASNGCV 240
 DB 181 QONGIGVYLNASTCPKPDPIPSHFLRVVNDSPCEKILPMDKSVDFLEKAKASNGCV 240
 QY 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKN 300
 DB 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKN 300
 QY 301 QTGASGPKSKLKLHLEKNEPVAVSEGGQKSTPLSPCADSATSEAGORPVHPASV 360
 DB 301 QTGASGPKSKLKLHLEKNEPVAVSEGGQKSTPLSPCADSATSEAGORPVHPASV 360
 QY 361 PSVPVQPSLLEDSPLVQALSGHLASADRLSDNSKLRSPSLDIKSVYSASMAASLHGF 420

```

Db      361 PSVPSVQPSLLEDSPLVQALSGHLHLSADRLSDSNKLSFSLDIKSVSYASMAASLHG
QY      421 SSEDALLEYKPTTLDGNTKLCQFSPVOELSEQTPETSPPDKEBASIPKKLQTARPSDQ
Db      421 SSEDALLEYKPTTLDGNTKLCQFSPVOELSEQTPETSPPDKEBASIPKKLQTARPSDQ
QY      481 SKRLHSVRTSSSGTAAQRLSLPLHRSQGYEDNYHTSFLGLSTSOQHLTKSAGLGLKGMH
Db      481 SKRLHSVRTSSSGTAAQRLSLPLHRSQGYEDNYHTSFLGLSTSOQHLTKSAGLGLKGMH
QY      541 SDILAPOSTPLTSSWYFATESSHFYASATYGGASATYASCSQLPCTGQVYSVRR
Db      541 SDILAPOSTPLTSSWYFATESSHFYASATYGGASATYASCSQLPCTGQVYSVRR
QY      601 QKPSDRADRRSRWHESSPEKQFKRRSCOMFEGESIMSENREELGKVGQSOSFSGSME
Db      601 QKPSDRADRRSRWHESSPEKQFKRRSCOMFEGESIMSENREELGKVGQSOSFSGSME
QY      661 IIEVS 665
Db      661 IIEVS 665

```

RESULT 14

ABR52352

ID ABR52352 standard; protein; 665 AA.

ABR52352;

19-JUN-2003 (first entry)

Protein relating to the invention SEQ ID NO: 42.

antiapoptotic; hepatotropic; nephrotropic; antiarrhythmic;
 antiproliferative; cardiatic; cytoskeletal; gene therapy; liver disease;
 immunological disorder; renal failure; cardiovascular disorder;
 immunological disorder; arthritis; psoriasis; congenital heart defect;
 congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

Homo sapiens.

WO200257460-A2.

25-JUN-2002.

20-DEC-2001; 2001WO-US050459.

20-DEC-2001; 2000US-0256868P.

30-MAR-2001; 2001US-0280186P.

01-MAY-2001; 2001US-0287735P.

05-JUN-2001; 2001US-0295848P.

25-JUN-2001; 2001US-0300465P.

(BRIM) BRISTOL-MYERS SQUIBB CO.

JACKSON DG, Feder J, Nelson T, Mintler G, Ramanathan C, Lee L,

Siemens N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,

Krystek S, Mcatee P, Suchard S, Banas D;

WPI; 2002-599721/64.

N-PSDB; ACC60521.

Novel polynucleotides encoding human phosphatase polypeptides useful in

the prevention or treatment of e.g. proliferative and cardiovascular

disorders.

Claim 5; Fig 5; 801pp; English.

The invention relates to a novel isolated nucleic acid comprising a

polynucleotide having a nucleotide sequence selected from 40

polynucleotides fully defined in the specification. The polynucleotide of

the invention has antiapoptotic, hepatotropic, nephrotropic,

antiarrhythmic, antiproliferative, cardiatic, and cytoskeletal activity. The
 polynucleotide may have a use in gene therapy. A polynucleotide or
 polypeptide of the invention is useful for preventing, treating or
 ameliorating a medical condition, e.g. a proliferative disorder. They are
 also useful for treating e.g. liver disease, renal failure, immunological
 disorders including arthritis and psoriasis, cardiovascular disorders
 such as congenital heart defects and congestive heart failure, and
 cancer. A method of the invention is useful for diagnosing a pathological
 condition or susceptibility to a condition in a subject. The present
 sequence is used in the exemplification of the invention

Sequence 665 AA;

Query Match 99.1%; Score 3388; DB 5; Length 665;

Best Local Similarity 99.4%; Pred. No. 6 7e-255; Matches 661; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSSTGTEKYLIDSRPFVYNTSHILEAININCKLMKRLQ 60
 Db 1 MAHEMIGTQIVTERLVALLSSTGTEKYLIDSRPFVYNTSHILEAININCKLMKRLQ 60

QY 61 DKVLTIELIQAHSKHKVDIDCQKVVVYDQSSQDVASISDCLTVLLGLKESFNSVHL 120
 Db 61 DKVLTIELIQAHSKHKVDIDCQKVVVYDQSSQDVASISDCLTVLLGLKESFNSVHL 120

QY 121 LAGGFAEFRCFPGLECKSTLVPTCISQCLPVANIGFTRILPNLYLGCQDVLNKEIL 180
 Db 121 LAGGFAEFRCFPGLECKSTLVPTCISQCLPVANIGFTRILPNLYLGCQDVLNKEIL 180

QY 181 QQNGIGVYLAATYTCRPFIPESHFLRVVNDSCFCKILPWIDKVDPIERAKASNGCV 240
 Db 181 QQNGIGVYLAATYTCRPFIPESHFLRVVNDSCFCKILPWIDKVDPIERAKASNGCV 240

QY 241 LVHCLAGISRSATIAIYIMKMDSLDAVYFVEKERTIPNPFGLDILDEKIKN 300
 Db 241 LVHCLAGISRSATIAIYIMKMDSLDAVYFVEKERTIPNPFGLDILDEKIKN 300

QY 301 QTGASGPKSKLKLHLEKNEBPVAVSEGGOKSETPSPCADSATSEMAQORPVHPASV 360
 Db 301 QTGASGPKSKLKLHLEKNEBPVAVSEGGOKSETPSPCADSATSEMAQORPVHPASV 360

QY 361 PSVPSVQPSLLEDSPLVQALSGHLHLSADRLSDSNKLSFSLDIKSVSYASMAASLHG 420
 Db 361 PSVPSVQPSLLEDSPLVQALSGHLHLSADRLSDSNKLSFSLDIKSVSYASMAASLHG 420

QY 421 SSEDALLEYKPTTLDGNTKLCQFSPVOELSEQTPETSPPDKEBASIPKKLQTARPSDQ 480
 Db 421 SSEDALLEYKPTTLDGNTKLCQFSPVOELSEQTPETSPPDKEBASIPKKLQTARPSDQ 480

QY 481 SKRLHSVRTSSSGTAAQRLSLPLHRSQGYEDNYHTSFLGLSTSOQHLTKSAGLGLKGMH 540
 Db 481 SKRLHSVRTSSSGTAAQRLSLPLHRSQGYEDNYHTSFLGLSTSOQHLTKSAGLGLKGMH 540

QY 541 SDILAPOSTPLTSSWYFATESSHFYASATYGGASATYASCSQLPCTGQVYSVRR 600
 Db 541 SDILAPOSTPLTSSWYFATESSHFYASATYGGASATYASCSQLPCTGQVYSVRR 600

QY 601 QKPSDRADRRSRWHESSPEKQFKRRSCOMFEGESIMSENREELGKVGQSOSFSGSME 660
 Db 601 QKPSDRADRRSRWHESSPEKQFKRRSCOMFEGESIMSENREELGKVGQSOSFSGSME 660

QY 661 IIEVS 665
 Db 661 IIEVS 665

RESULT 15

ABR52424

ID ABR52424 standard; protein; 664 AA.

ABR52424;

19-JUN-2003 (first entry)

XX Protein relating to the invention SEQ ID NO: 190.

XX anti-proliferative; hepatotropic; nephrotropic; anti-arthritic;
 XX anti-proliferative; cardiatic; cytostatic; gene therapy; liver disease;
 XX proliferative disorder; renal failure; cardiovascular disorder;
 XX immunological disorder; arthritis; psoriasis; congenital heart defect;
 XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

OS Homo sapiens.

XX WO200257460-A2.

XX 25-JUL-2002.

XX 20-DEC-2001; 2001WO-US050459.

XX 20-DEC-2000; 2000US-0256868P.

XX 30-MAR-2001; 2001US-0280186P.

XX 01-MAY-2001; 2001US-0287735P.

XX 25-JUN-2001; 2001US-0295848P.

XX 25-JUN-2001; 2001US-0300465P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Jackson DQ, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;

XX Stiemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;

XX Krysiek S, Mcatee P, Suchard S, Banas D;

XX WPI; 2002-599721/64.

XX Novel polynucleotides encoding human phosphatase polypeptides useful in

XX the prevention or treatment of e.g. proliferative and cardiovascular

XX disorders.

XX Example 57; Page 500-501; 801pp; English.

XX The invention relates to a novel isolated nucleic acid comprising a

XX polynucleotide having a nucleotide sequence selected from 40

XX polynucleotides fully defined in the specification. The polynucleotide of

XX the invention has anti-proliferative, hepatotropic, nephrotropic,

XX anti-arthritic, anti-proliferative, cardiatic, and cytostatic activity. The

XX polynucleotide may have a use in gene therapy. A polynucleotide or

XX polypeptide of the invention is useful for preventing, treating or

XX ameliorating a medical condition, e.g. a proliferative disorder. They are

XX also useful for treating e.g. liver disease, renal failure, immunological

XX disorders including arthritis and psoriasis, cardiovascular disorders

XX such as congenital heart defects and congestive heart failure, and

XX cancer. A method of the invention is useful for diagnosing a pathological

XX condition or susceptibility to a condition in a subject. The present

XX sequence is used in the exemplification of the invention

XX Sequence 664 AA;

Query Match 98.6%; Score 3368.5; DB 5; Length 664;

Best Local Similarity 99.1%; Pred. No. 2.2e-253;

Matches 659; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 MAHEMIGTQVTEVTEVLLVLESGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
 DB 1 MAHE-IGTQVTEVTEVLLVLESGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 59
 QY 61 DKVLITELIQHSAKGVYDIDCSQKVVYDQSSQDVASLSDCLTVLLGKLEKSFNSVHL 120
 DB 60 DKVLITELIQHSAKGVYDIDCSQKVVYDQSSQDVASLSDCLTVLLGKLEKSFNSVHL 119
 QY 121 LAGGFAEFSRCFPGICGKSTLVPTCTSQCLPVAINGPRTIIPNLVYGCQDVLANKELEI 180
 DB 120 LAGGFAEFSRCFPGICGKSTLVPTCTSQCLPVAINGPRTIIPNLVYGCQDVLANKELEI 179
 QY 181 QONGIGVYLNASVYCPKPDFIPESHFLRPVNDSECEKILPWLIDKSVDFIEKAKANGCV 240
 DB 180 QONGIGVYLNASVYCPKPDFIPESHFLRPVNDSECEKILPWLIDKSVDFIEKAKANGCV 239

QY 241 LVHCLAGISRSATIAIAYIMKRWMSLDEAYRPFVKEKRPITISPNFPLGOLLPEKKIKN 300
 DB 240 LVHCLAGISRSATIAIAYIMKRWMSLDEAYRPFVKEKRPITISPNFPLGOLLPEKKIKN 299
 QY 301 QGASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPLSPPCADSATSBAAGRPVHPASV 360
 DB 300 QGASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPLSPPCADSATSBAAGRPVHPASV 359
 QY 361 PSVPSVQPSLLEDSPVQALSGHLSDADRLSDSNKLKRSFLDIKVSYSASMAASLHGF 420
 DB 360 PSVPSVQPSLLEDSPVQALSGHLSDADRLSDSNKLKRSFLDIKVSYSASMAASLHGF 419
 QY 421 SSEDALFYKPPSTTLDGNTKLCQPSPVQELSPQTEPSPDXEASIPKLTOTARPESDQ 480
 DB 420 SSEDALFYKPPSTTLDGNTKLCQPSPVQELSPQTEPSPDXEASIPKLTOTARPESDQ 479
 QY 481 SKRLHSVRTSSSGTAQRSLSPHRSQSVEDNYHTSFLFGLSTSOQHLTKSAGLGLKGMH 540
 DB 480 SKRLHSVRTSSSGTAQRSLSPHRSQSVEDNYHTSFLFGLSTSOQHLTKSAGLGLKGMH 539
 QY 541 SDILAPQSTPSLTSSWYPATESSHFYASASAIYGSASYSAYSCQLPTCGDOVYSVRRR 600
 DB 540 SDILAPQSTPSLTSSWYPATESSHFYASASAIYGSASYSAYSCQLPTCGDOVYSVRRR 599
 QY 601 QKESDADSRKSWHESPPKOPKRRSCOMEFESIMSENRSPREELGXVGSOSFSGSME 660
 DB 600 QKESDADSRKSWHESPPKOPKRRSCOMEFESIMSENRSPREELGXVGSOSFSGSME 659
 QY 661 IIEVS 665
 DB 660 IIEVS 664

Search completed: June 21, 2004, 13:20:33
 Job time : 76.2709 sec

QY	481	SKRLHSVFRSSSGTARSLSLPLHREGSVEDNDHTHTFLPGLSTNSQOHLTYSAGLGLKGMH	540
Db	476	VTRLHSVRKSGSSSTPRPFSSPLHREGSVEDNDHTHTFLPGLSTNSQOHLTYSAGLGLKGMH	535
QY	541	SDILAPQTSPLTSSWYFATESSHFYASAIYGSGASYSAYSCSOLPTCGDOVYSVRR	600
Db	536	SDILAPQASAPLTSWYFATEPSHLYASAIYGNSSYAYSCGQLPTCSDDIYVRR	595
QY	601	QKPSRADRRSRWHESSPPEKQPKRRSCOMEFEESIMSENRSREELGRVSGSFGSGME	660
Db	556	QKPTPRADRRSRWHESSPPEKQPKRRSCOMEFEESIMSENRSREELGRVSGSFGSGME	655
QY	661	IIEVS 665	
Db	656	IIEVS 660	
RESULT 2			
ID	AAAB20325		
XX	AAAB20325 standard; protein; 666 AA.		
XX	AAAB20325;		
AC	AAAB20325;		
DT	29-MAY-2001 (first entry)		
XX			
DE	Human protein phosphatase and kinase protein-4.		
XX			
KW	Protein phosphatase and kinase protein; PPKR-4; human;		
KW	gastrointestinal disorder; immune system disorder; neurological disorder;		
KW	cell proliferative disorder; cancer; diagnosis; therapy.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FH	Modified-site	12	
FT		/note= "O-phosphorylated"	
FT	Modified-site	21	
FT		/note= "O-phosphorylated"	
FT	Modified-site	23	
FT		/note= "O-phosphorylated"	
FT	Modified-site	38	
FT		/note= "N-glycosylated"	
FT	Modified-site	38	
FT		/note= "N-glycosylated"	
FT	Modified-site	49	
FT		/note= "N-glycosylated"	
FT	Modified-site	72	
FT		/note= "O-phosphorylated"	
FT	Modified-site	82	
FT		/note= "O-phosphorylated"	
FT	Region	85..298	
FT		/note= "VH1-type dual specificity phosphatase signature"	
FT	Modified-site	91	
FT		/note= "O-phosphorylated"	
FT	Modified-site	190	
FT		/note= "N-glycosylated"	
FT	Modified-site	212	
FT		/note= "N-glycosylated"	
FT	Modified-site	214	
FT		/note= "O-phosphorylated"	
FT	Active-site	220..280	
FT		/note= "tyrosine specific protein phosphatase"	
FT	Region	237..278	
FT		/note= "Y phosphatase signature"	
FT	Modified-site	266	
FT		/note= "O-phosphorylated"	
FT	Modified-site	280	
FT		/note= "O-phosphorylated"	
FT	Modified-site	300	
FT		/note= "N-glycosylated"	
FT	Modified-site	369	
FT		/note= "O-phosphorylated"	
FT	Modified-site	393	
FT		/note= "O-phosphorylated"	

FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	421
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	422
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	434
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	439
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	468
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	471
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	479
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	528
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	590
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	597
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	605
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	610
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	613
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	618
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	628
FT	Modified-site	/note= "O-phosphorylated"
FT	Modified-site	641
FT	Modified-site	/note= "N-glycosylated"
FT	Modified-site	643
FT	Modified-site	/note= "O-phosphorylated"
PN	WO200120004-A2.	
XX		
PD	22-MAR-2001.	
XX		
PE	14-SEP-2000; 2000MO-US025515.	
XX		
PR	15-SEP-1999; 99US-0154141P.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
PI	Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y;	
PI	Lu DM;	
DR	WPI; 2001-244811/25.	
XX		
DR	N-PSDB; AAF30479.	
XX		
PT	Novel human protein phosphatase and kinase proteins for diagnosis,	
PT	treatment and prevention of gastrointestinal, immune system, neurological	
PT	and cell proliferative disorders.	
XX		
PS	Claim 1; Page 87-88; 103pp; English.	
XX		
CC	The present sequence is that of novel human protein phosphatase and	
CC	kinase protein PPHK-4, as predicted from Incyte Clone ID No. 1234795CB1	
CC	(see AAF30479). Tissues that express PPHK-4 (as a fraction of total	
CC	tissues expressing PPHK-4) include gastrointestinal (0.385),	
CC	cardiovascular (0.154), nervous (0.154) and reproductive (0.154).	
CC	Diseases or conditions associated with tissues expressing PPHK-4 (as a	
CC	fraction of total tissues expressing PPHK-4) include cancer (0.692),	
CC	inflammation or trauma (0.308) and cell proliferation (0.231). The	
CC	encoded protein shows homology to mouse neuronal tyrosine threonine	
CC	phosphatase 1. The invention provides human PPHK-1 to -11 polypeptides	
CC	(see AAF2032-32) and polynucleotides (see AAF30476-86). It also provides	
CC	expression vectors, host cells, antibodies, agonists and antagonists, as	
CC	well as methods for diagnosing, treating or preventing disorders	
CC	associated with expression of PPHK, including gastrointestinal	
CC	disorders, immune system disorders, neurological disorders and cell	

CC proliferative disorders, including cancer
 XX Sequence 666 AA;
 Query Match 88.7%; Score 3030.5; DB 1; Length 666;
 Best Local Similarity 88.9%; Pred. No. 0;
 Matches 607; Conservative 8; Mismatches 33; Indels 35; Gaps 5;

QY 1 MAHEMIGTOIVTERLVALESSTGTEKVLIDSRPFVEVNTSHLEAINININSKLMKRLQ 60
 DB 1 MAHEMIGTOIVTERLVALESSTGTEKVLIDSRPFVEVNTSHLEAINININSKLMKRLQ 60
 QY 61 DKVLITELIOWSAKHVVIDCSQKVVVYDQSSQDVASLSDDCEFLVYLGLKLEKSPNSVHL 120
 DB 61 DKVLITELIOWSAKHVVIDCSQKVVVYDQSSQDVASLSDDCEFLVYLGLKLEKSPNSVHL 120
 QY 121 LAGGFABFRCFPGLECKSTLVPCTISQPCLPVANIIGPRLIPNLVYLGCQDVLNKEIM 180
 DB 121 LAGGFABFRCFPGLECKSTLVPCTISQPCLPVANIIGPRLIPNLVYLGCQDVLNKEIM 180
 QY 181 QONGIGVLNASTCPKDPFIPESHFLRVVNDSPCEKILPMLDKSVDFTEKAKASNGCV 240
 DB 181 QONGIGVLNASTCPKDPFIPESHFLRVVNDSPCEKILPMLDKSVDFTEKAKASNGCV 240
 QY 241 LVHCLAGISRATIAIAYIKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKIKIN 300
 DB 241 LVHCLAGISRATIAIAYIKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKIKIN 300
 QY 301 QTGASGPKSLKLIHLEKNEPVPAVSEGGQKSETPISPCADSAISEAAGRPVHPASV 360
 DB 301 QTGASGPKSLKLIHLEKNEPVPAVSEGGQKSETPISPCADSAISEAAGRPVHPASV 360
 QY 361 PSVPVOPSLIEDSPLVQALSGHLASDRLEDSNKLKRSPLDIKSVSASMAASLHGF 420
 DB 361 PSVPVOPSLIEDSPLVQALSGHLASDRLEDSNKLKRSPLDIKSVSASMAASLHGF 420
 QY 421 SSEDLELVKKPSTLTDGNTKLCQFSPVQELSECTPETS-----PDKEBASIPKQLQTA 474
 DB 421 SSEDLELVKKPSTLTDGNTKLCQFSPVQELSECTPETS-----PDKEBASIPKQLQTA 474
 QY 475 RPDSOSKR-----LHSVTSSTSGTQORSLSLPHNSGVEDNYHTSFLGLS 522
 DB 475 RPDSOSKR-----LHSVTSSTSGTQORSLSLPHNSGVEDNYHTSFLGLS 522
 QY 523 TSQOHLTKSAGLGLKGMHSDIILAPOTSTPLTSWYATSSHPYASATYGGASASAY 582
 DB 523 TSQOHLTKSAGLGLKGMHSDIILAPOTSTPLTSWYATSSHPYASATYGGASASAY 582
 QY 583 SCQOLPTCGDQVYSVRRRQKPSDRADSRSMHESPEKQFKRRSCQMEFGESIMENRS 642
 DB 583 SCQOLPTCGDQVYSVRRRQKPSDRADSRSMHESPEKQFKRRSCQMEFGESIMENRS 642
 QY 643 REELGVKVSQSPFSGSWEIIEVS 665
 DB 643 REELGVKVSQSPFSGSWEIIEVS 665
 QY 644 REELGVKVSQSPFSGSWEIIEVS 666
 DB 644 REELGVKVSQSPFSGSWEIIEVS 666

RESULT 3
 AAM25744
 ID AAM25744 standard; protein; 672 AA.
 XX
 AC AAM25744;
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:1259.
 XX
 XX Human; cancer; HIV infection; human immunodeficiency virus;
 KM anti-inflamatory; anti-rheumatic; anti-arthritic; immunosuppressive;
 KM antibacterial; endocrine; cardiac; central nervous system; virocid;
 KM anti-HIV; fungicide; antitumor; cardiovascular; antiaesthetic; anemol;
 KM antiagregant; haemostatic; vulnary; antileuk; osteopathic; eczema;
 KM dermatological; antiallergic; antiaesthetic; antidiabetic; cyostatic;

KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KM antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KM genetic diseases; haematopoietic disorder; platelet disorder; asthma;
 KM chromocytopenia; osteoporosis; severe combined immunodeficiency;
 KM allergic rhinitis; diabetes; multiple sclerosis; depression;
 KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KM neurological disorder.
 XX
 OS Homo sapiens.
 PN WO200153455-A2.
 PD 26-JUL-2001.
 XX
 PF 22-DEC-2000; 2000MO-US035017.
 XX
 XX 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-457603/49.
 DR N-PSDB; AAM99685.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
 PS Claim 20; Page 260; 1217pp; English.
 XX
 CC AAM99166 to AAM99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and cells
 CC they are expressed in, such as: anti-inflamatory; antirheumatic;
 CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiac;
 CC central nervous system; virocid; anti-HIV; fungicide; antitumor;
 CC cardiovascular; antiaesthetic; antiagregant; haemostatic; vulnary;
 CC antidiabetic; osteopathic; dermatological; antiallergic; antiaesthetic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders
 XX
 SO Sequence 672 AA;
 Query Match 85.7%; Score 2930; DB 1; Length 672;
 Best Local Similarity 86.5%; Pred. No. 0;
 Matches 591; Conservative 10; Mismatches 46; Indels 36; Gaps 5;

QY 1 MAHEMIGTOIVTERLVALESSTGTEKVLIDSRPFVEVNTSHLEAINININSKLMKRLQ 60
 DB 8 MAHEMIGTOIVTERLVALESSTGTEKVLIDSRPFVEVNTSHLEAINININSKLMKRLQ 67
 QY 61 DKVLITELIOWSAKHVVIDCSQKVVVYDQSSQDVASLSDDCEFLVYLGLKLEKSPNSVHL 120
 DB 61 DKVLITELIOWSAKHVVIDCSQKVVVYDQSSQDVASLSDDCEFLVYLGLKLEKSPNSVHL 120
 QY 121 LAGGFABFRCFPGLECKSTLVPCTISQPCLPVANIIGPRLIPNLVYLGCQDVLNKEIM 180
 DB 121 LAGGFABFRCFPGLECKSTLVPCTISQPCLPVANIIGPRLIPNLVYLGCQDVLNKEIM 180
 QY 181 QONGIGVLNASTCPKDPFIPESHFLRVVNDSPCEKILPMLDKSVDFTEKAKASNGCV 240
 DB 181 QONGIGVLNASTCPKDPFIPESHFLRVVNDSPCEKILPMLDKSVDFTEKAKASNGCV 240
 QY 241 LVHCLAGISRATIAIAYIKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKIKIN 300
 DB 241 LVHCLAGISRATIAIAYIKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKIKIN 300
 QY 301 QTGASGPKSLKLIHLEKNEPVPAVSEGGQKSETPISPCADSAISEAAGRPVHPASV 360
 DB 301 QTGASGPKSLKLIHLEKNEPVPAVSEGGQKSETPISPCADSAISEAAGRPVHPASV 360
 QY 361 PSVPVOPSLIEDSPLVQALSGHLASDRLEDSNKLKRSPLDIKSVSASMAASLHGF 420
 DB 361 PSVPVOPSLIEDSPLVQALSGHLASDRLEDSNKLKRSPLDIKSVSASMAASLHGF 420
 QY 421 SSEDLELVKKPSTLTDGNTKLCQFSPVQELSECTPETS-----PDKEBASIPKQLQTA 474
 DB 421 SSEDLELVKKPSTLTDGNTKLCQFSPVQELSECTPETS-----PDKEBASIPKQLQTA 474
 QY 475 RPDSOSKR-----LHSVTSSTSGTQORSLSLPHNSGVEDNYHTSFLGLS 522
 DB 475 RPDSOSKR-----LHSVTSSTSGTQORSLSLPHNSGVEDNYHTSFLGLS 522
 QY 523 TSQOHLTKSAGLGLKGMHSDIILAPOTSTPLTSWYATSSHPYASATYGGASASAY 582
 DB 523 TSQOHLTKSAGLGLKGMHSDIILAPOTSTPLTSWYATSSHPYASATYGGASASAY 582
 QY 583 SCQOLPTCGDQVYSVRRRQKPSDRADSRSMHESPEKQFKRRSCQMEFGESIMENRS 642
 DB 583 SCQOLPTCGDQVYSVRRRQKPSDRADSRSMHESPEKQFKRRSCQMEFGESIMENRS 642
 QY 643 REELGVKVSQSPFSGSWEIIEVS 665
 DB 643 REELGVKVSQSPFSGSWEIIEVS 665
 QY 644 REELGVKVSQSPFSGSWEIIEVS 666
 DB 644 REELGVKVSQSPFSGSWEIIEVS 666

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:15:52 ; Search time 22.6939 Seconds

(Without alignments)
2818.703 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 3418

Sequence: 1 MAHEMGTGIVTERLVALLR.....LKNVSGSSPSGSMELIIVS 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	497.5	14.6	619	2	T15969
2	443	13.0	394	2	A56115
3	433	12.7	367	1	S24411
4	423	12.4	367	2	S52265
5	420	12.3	367	1	S29090
6	418	12.2	314	1	A57126
7	409	12.0	314	2	B57126
8	393.5	11.5	393	2	A56947
9	360.5	10.5	384	1	I38890
10	339	9.9	303	1	T46405
11	294.5	8.6	365	2	T32494
12	239.5	7.0	272	2	T18915
13	233.5	6.8	226	2	T21380
14	225.5	6.6	186	2	T16056
15	223.5	6.5	330	2	T39698
16	208.5	6.1	220	2	JC7885
17	205.5	6.0	866	2	F88481
18	205	6.0	223	2	I49365
19	204	6.0	278	2	T39517
20	203.5	6.0	771	2	T47666
21	202	5.9	142	2	T03074
22	202	5.9	364	1	S31304
23	195	5.7	185	1	A47196
24	195	5.7	283	2	G84458
25	193	5.6	205	2	I49364
26	189	5.5	807	1	S58755
27	185	5.4	276	2	T48906
28	183.5	5.4	204	2	T17802
29	178	5.2	204	2	T17802

30	177	5.2	580	2	T18439	hypothetical prote
31	172	5.0	600	2	T18446	hypothetical prote
32	170.5	5.0	282	2	S41012	hypothetical prote
33	157	4.6	209	1	S48459	probable dual spec
34	150.5	4.4	169	2	T30684	probable dual spec
35	147.5	4.3	597	1	S43743	probable dual spec
36	146.5	4.3	272	2	T19418	hypothetical prote
37	145.5	4.3	171	1	T36845	dual specificity p
38	145.5	4.3	171	2	T28522	probable dual spec
39	145.5	4.3	171	2	B72161	dual protein - vari
40	144.5	4.2	171	1	QOVZHI	dual specificity p
41	142.5	4.2	928	2	S50578	hypothetical prote
42	140	4.1	2271	2	F90073	hypothetical prote
43	139.5	4.1	534	2	T39903	serine-rich protei
44	138.5	4.1	171	1	A42514	dual specificity p
45	137.5	4.0	1906	1	S68235	myosin-light-chain

ALIGNMENTS

RESULT 1

T15969

hypothetical protein F08B1.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

R/Chisoe, S.

submitted to the EMBL Data Library, July 1995

A/Description: The sequence of C. elegans cosmid F08B1.

A/Reference number: Z16439

A/Accession: T15969

A/Status: preliminary; translated from GB/EMBL/DDBT

A/Molecule type: DNA

A/Residues: 1-619 <CHI>

A/Cross-references: EMBL:U23178; NID:9726421; PID:9726422; PIDN:AAC6719.1; CESP:F08B1.1

C/Genetics:

A/Experimental source: strain Bristol N2

A/Gene: CESP:F08B1.1

A/Intons: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Query Match	14.6%	Score 497.5;	DB 2;	Length 619;
Best Local Similarity	28.9%	Pred. No. 4.4e-24;		
Matches	160;	Conservative	77;	Mismatches 159; Indels 157; Gaps 19;
QY	123	GGFAERSRCFGLCEGKS--TLVPTISQPL--PVANIGPTNLPMLYGCQRDVIANKE	178	
DB	99	GGFKQPAQYPCQCSSESGWTRLPQSLQPLQPTGD-GITLITPVIYGSQIDSIDET	157	
QY	179	LIQNGIGYLVANSYTCPRKDFIPE-SHPLRVVNVNSFCCKILPWLKDSVDFTEKAKASN	237	
DB	158	MLDALDISVIVINISMTCPSKVCIKEDKFNIRIVNDSYQKELDPYFMAVEFEKCRAG	217	
QY	238	GCVLVHCLAGISRSATIAIYIMKRMDSIDEAYRFYKERTPTISPFNFILGQLDYEK-	296	
DB	218	KKCLIHCLAGISRSATIAIYIMKRMDSIDEAYRFYKERTPTISPFNFILGQLDYENV	277	
QY	297	---KIKNQTAGSGPKSKLKL--HLEKNEP--	329	
DB	278	LKHQVLDVYQAQRPHRMDYVGPDLCPKPVKASNSNCVPGSTHDESSPSPVSEG	337	
QY	330	GQSEETPLSPCADSATSEAGORPVHPASVPSVQPSLDBSDPLVQALSGHLISADR	389	
DB	338	SAASEETSSAASSSTASA-----PPSPSTSE-----QGTSSGTVAVNG	379	
QY	390	LEDSNKLKSGFSLDI-----KSVYSASMAALHGFSSSEBDALVYKPSFTLLDGTNKL	443	
DB	380	-----KKNMTMDLGLPHRPAKALGLPSRIGTSV-----AELPSFTLSRLS---	420	
QY	444	QSPVQVLESGTP--ETSPDKKEASIPKQLQTPRPSDSQSKLHVSRTSSGTAQASLL	500	
DB	421	-FNGPEKIAVSTPIILNFTNCPNSPIIPV-----ASSSRREVILTLTP	460	

QY 501 SPLHRSQSVEDNHTSFLFGLSTSOOHLTKSAGLGLKGMHSDILAPQTSPELSTSSWYFA 560
| | | | |
Db TFAASSSS-----STS-----SEPFDDSSPES 483
QY 561 TESSH-----FYSAATYGGSAYSYAVSCQLPTCGQVYSVRBRQRKPSDRADRRSMW 614
| | | | |
Db 484 SSSSSIVENPFPASTEVPAGSSSISTSPGSG-----STPASASSAARSC 529
QY 615 EESPEKOPKRRS 627
| | | | |
Db 530 RMKGFFKVFPSKKA 542

RESULT 2

A56115

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 4 - human

N/Alternate names: dual specificity phosphatase HVH2

C/Species: Homo sapiens (man)

C/Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Apr-1998

C/Accession: A56115

R/Guan, K.L.; Butch, E.

J. Biol. Chem. 270, 7197-7203, 1995

A/Title: Isolation and characterization of a novel dual specific phosphatase, HVH2, which

A/Reference number: A56115; MUID:95221370; PMID:7535768

A/Accession: A56115

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-394 <GUA>

A/Cross-references: GB:U21108

C/Genetics:

A/Gene: GDB:DUSP4; HVH2; MKP-2

A/Cross-references: GDB:433893

A/Map position: 8p21-8p11.2

C/Suprafamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity

C/Keywords: phosphoprotein; phosphoric monoester hydrolase

F/203-334/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

F/280/Active site: Cys (phosphocysteine intermediate) #status predicted

F/286/Binding site: substrate phosphate (Arg) #status predicted

Query Match 13.0%; Score 443; DB 2; Length 394;
Best Local Similarity 31.4%; Pred. No. 6.9e-21;
Matches 120; Conservative 74; Mismatches 142; Indels 46; Gaps 14;

QY 19 LESGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQODKVLITELI--CHSAKH 76
| | | | |
Db 39 LPSSG-GKCLIDCPFLAHSAGYILGYSVNVRCNTIVRR-AGSVSLQGLIPAEVRRAR 96
QY 77 VDIDCSQKVYVYDSSQDVASLSDCFITVLIGLKSF--NSVHLLAGGFAEFSRCPPG 134
| | | | |
Db 97 LRSGLYSAVIVYDERSPPAESLRBDSVSLVVOALRRNAERTDCLCKGGERFSSEYPE 156
QY 135 LCEGKSTL-----VPTCISOP-----C-LPVANI-GPTRILPNIYLGGQRDVNLKEL 179
| | | | |
Db 157 FCSGTXKMLAATPPVPBATEPDLGSSCCTPLHDGGPVEILLFYLIGSAHYAARDM 216
QY 180 IQONGIGVYVNASYTCPKPDPFIPESHFLRPVNDSPCEKILPWLDSVDFIEKAKANGC 239
| | | | |
Db 217 LDALGITALINVSDDCPN-HFEGHYQKCIPEVDNHRADISSWMEALIEYIDAVKDCGR 275
QY 240 VLVHCLAGISRSATITAYIMKRMDSIDEARFVKERPTISNPNFLGQLDYEKKIK 299
| | | | |
Db 276 VLVHCGAGISRSATITAYIMKRMKRVLEAEFVKORRSTISNPFSGQLQFESQVL 335
QY 300 NOTGASGPKSKLKLHLEKNEPVPVAVSEGQKSETPPCADATSEAAAGRPVHPAS 359
| | | | |
Db 336 ATSCAARAS-----PSGPL-----REKKTATP-----TSQVFSPFPV-SVG 373
QY 360 VPSVPSVQPSLLEDSPLVQALS 381
| | | | |
Db 374 VHSAPSSLPYL--HSPITTPS 393

RESULT 3

S24411

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - mouse

N/Alternate names: 3CH134 protein; protein-tyrosine-phosphatase erp, nonreceptor type 10

C/Species: Mus musculus (house mouse)

C/Date: 19-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999

C/Accession: A54681; S24411

R/Noguichi, T.; Metz, R.; Chen, L.; Mattei, M.G.; Carrasco, D.; Bravo, R.

Mol. Cell. Biol. 13, 5195-5205, 1993

A/Title: Structure, mapping, and expression of erp, a growth factor-inducible gene encodi

A/Reference number: A54681; MUID:93360956; PMID:8355678

A/Accession: A54681

A/Molecule type: DNA

A/Residues: 1-367 <NOG>

A/Cross-references: GB:S64851; NID:g409976; PIDN:AA27882.1; PID:g409977

R/Charles, C.H.; Adler, A.S.; Lau, L.F.

Oncogene 7, 187-190, 1992

A/Title: cDNA sequence of a growth factor-inducible immediate early gene and characteriz

A/Reference number: S24411; MUID:92158357; PMID:1741163

A/Accession: S24411

A/Molecule type: mRNA

A/Residues: 1-367 <CHA>

A/Cross-references: EMBL:X61940; NID:g49735; PIDN:CAA43944.1; PID:g49736

C/Genetics:

A/Gene: erp

A/Introns: 123/1; 172/1; 245/1

C/Suprafamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity

C/Keywords: immediate-early protein; phosphoprotein; phosphoric monoester hydrolase

F/181-312/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

F/258/Active site: Cys (phosphocysteine intermediate) #status predicted

F/264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.7%; Score 433; DB 1; Length 367;
Best Local Similarity 31.6%; Pred. No. 2.7e-20;
Matches 115; Conservative 54; Mismatches 151; Indels 44; Gaps 8;

QY 15 LVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQODKVLITELIOLHSAR 74
| | | | |
Db 13 LRALLREGAQCLLDRCSPFAFNAGHIGSVVNRSTIVRRAKGAMGLEHIVPAEIR 72
QY 75 HKVDIDCSQKVYVYDSSQDVASLSDCFITVLIGL--EKSNSVHLLAGGFAEFSRCPP 132
| | | | |
Db 73 GRLLAGAYHVVLLDRSASLIDGAKRDTGLAALGCREARSTQVFFLGGGEARSASC 132
QY 133 PGLCEGKSTLVPTCISOP-----CLPVANI-----GPTRIIPNIYLGGQRDVNL 175
| | | | |
Db 133 PELCSQST--PGLSLPLSTVPDASGSCSCTPLPDGGPVEILLFYLIGSAHYAS 190
QY 176 NKELIQONGIGVYVNASYTCPKPDPFIPESHFLRPVNDSPCEKILPWLDSVDFIEKAKA 235
| | | | |
Db 191 RKDMLDALGITALLINVSANCPN-HFEGHYQKSIPEVDNHRADISSWMEALIDFIDSID 249
QY 236 SNGCVYVHCLAGISRSATITAYIMKRMDSIDEARFVKERPTISNPNFLGQLDYE 295
| | | | |
Db 250 AGGRVVFHCGAGISRSATITAYIMKRMKRVLEAEFVKORRSTISNPFSGQLQFESQVL 309
QY 296 KXIKNQTGASGPKSKLKLHLEKNEPVPVAVSEGQKSETPPCADATSEAAAGRPV 355
| | | | |
Db 310 SQV-----LAPHCS-----AEAGSPAAVLDKRGISTTVTFNFPVS-----ITV 347
QY 356 HPAS 359
| | | | |
Db 348 HPTN 351

RESULT 4

S52265

dual specificity phosphatase (EC 3.1.3.-) 1 - rat

N/Alternate names: protein-tyrosine-phosphatase CL100; protein-tyrosine-phosphatase, non

C/Species: Rattus norvegicus (Norway rat)

C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999

C/Accession: S52265

R/Muda, M.; Schlegel, W.; Arkinsteil, S.

submitted to the EMBL Data Library, January 1995

A/Description: Pathways regulating CL100 gene expression in pituitary cells.
 A/Reference number: S52265
 A/Accession: S52265
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-367 <MUD>
 A/Cross-references: EMBL:X64004; NID:G642264; PIDN:CA58828.1; PID:G642265
 C/Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
 C/Keywords: phosphoprotein, phosphoric monoester hydrolase
 F/181-312/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
 F/258/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.4%; Score 423; DB 2; Length 367;
 Best Local Similarity 31.3%; Pred. No. 1.2e-19;
 Matches 115; Conservative 54; Mismatches 148; Indels 50; Gaps 9;

Qy	15	LVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKMKRRLQ-----QDKVLTLELI 69
Db	13	LRALLGERRAQCILLDCRSFPAFNAGHTGVSVVNRSTIYRRKAKGMGLHIVPNAELR 71
Qy	70	QHSARKVNDIDCSQKVVVYDQSSQDVASLSDCFVLTKL--EKSFNSVHLAAGFAE 127
Db	72	-----RGRLLAGVAVVLLDERSAALDGAKRDGTLAAGALCREASTQVFLGGYE 127
Qy	128	FSRCFPGELCEGK-----STLVPTCISQPCLPVANI-----GPTRIPLNLYGCGR 172
Db	128	FSASCPELCSKQSTPMGLSLPLSTVSPDASGSCSTPLDQCGVLELFLYLSAY 187
Qy	173	DVANKELIQNGIGVYLNASTYCPKPDFIPESHFLRVVNDSPCEKILPWLKSVDFTEK 232
Db	188	HARKTMLDALGTTALINVSANCPN-HFEGHYQKSLPVEDNHKADISSPFNEAIDPIDS 246
Qy	233	AKASNGCVLVHCLAGISRSATTAIAYIMKMDSLDEAYRVYKERTISPNFNFLGQL 292
Db	247	IKDAGRGVFFHCAGISRSATICLAVLMRTNRVYLDFAEFVKQRRIISPFNSFMGQL 306
Qy	293	DYRKATIKNOTGASGPKSKLHLHLEKNEVPVASEGQKSET-----PLSPCADSATSEAGQ 352
Db	307	QFESQV-----LAFHCS-----AEAGSPMAVLDRTSTTVTFNFPVS----- 344
Qy	353	RPVHPAS 359
Db	345	IPVHPIN 351

RESULT 5
 dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human
 N/Alternate names: protein-tyrosine-phosphatase CL100; protein-tyrosine-phosphatase, not
 C/Species: Homo sapiens (man)
 C/Date: 25-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999
 C/Accession: S29090; A53052
 R/Keyes, S.M.; Emalle, E.A.
 Nature 359, 644-647, 1992
 A/Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine
 A/Reference number: S29090; MUID:93024952; PMID:1406996
 A/Accession: S29090
 A/Molecule type: mRNA
 A/Residues: 1-367 <KEY>
 A/Cross-references: EMBL:X68277; NID:G29980; PIDN:CA48338.1; PID:G29981
 R/Kwak, S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.
 J. Biol. Chem. 269, 3596-3604, 1994
 A/Title: Isolation and characterization of a human dual specificity protein-tyrosine pho
 A/Reference number: A53052; MUID:94148864; PMID:8106404
 A/Accession: A53052
 A/Molecule type: DNA
 A/Residues: 1-367 <KKA>
 A/Experimental source: leukocyte
 A/Note: Sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804,
 C/Genetics:
 A/Gene: GDB:DUSP1, PTPN10
 A/Cross-references: GDB:136197; OMIM:600714

A/Map position: 5q34-5q34
 C/Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
 C/Keywords: heat shock, phosphoprotein, phosphoric monoester hydrolase, stress-induced p
 F/181-312/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
 F/258/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.3%; Score 420; DB 1; Length 367;
 Best Local Similarity 31.2%; Pred. No. 1.8e-19;
 Matches 113; Conservative 56; Mismatches 159; Indels 34; Gaps 7;

Qy	15	LVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKMKRRLQDDKVLTELIOHSK 74
Db	13	LRALLGERRAQCILLDCRSFPAFNAGHTGVSVVNRSTIYRRKAKGMGLHIVPNAELR 72
Qy	75	HKYDIDCSQKVVVYDQSSQDVASLSDCFVLTKL--EKSFNSVHLAAGFAEFSRCF 132
Db	73	GRLLAGVAVVLLDERSAALDGAKRDGTLAAGALCREAARAOVFFLGGYEAFFSASC 132
Qy	133	PGICEBK-----STLVPTCISQPCLPVANI-----GPTRIPLNLYGCGRVLANK 177
Db	133	PELCSKQSTPMGLSLPLSTVSPDASGSCSTPLDQCGVLELFLYLSAYNASRK 192
Qy	178	ELIQNGIGVYLNASTYCPKPDFIPESHFLRVVNDSPCEKILPWLKSVDFTEKASN 237
Db	193	DMLDALGTTALINVSANCPN-HFEGHYQKSLPVEDNHKADISSWFNEAIDFIDSTINAG 251
Qy	238	GCYLVHCLAGISRSATTAIAYIMKMDSLDEAYRVYKERTISPNFNFLGQLDYEEK 297
Db	252	GRVFFVHCAGISRSATICLAVLMRTNRVYLDFAEFVKQRRIISPFNSFMGQLDQESQ 311
Qy	298	IKNOTGASGPKSKLHLHLEKNEVPVASEGQKSET-----PLSPCADSATSEAGQ 353
Db	312	V-----LAFHCS-----AEAGSPMAVLDRTSTTVTFNFPVSIPVHSTNSALSTLQS 359
Qy	354	PV 355
Db	360	PI 361

RESULT 6
 dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 2 - human
 N/Alternate names: mitogen-induced nuclear tyrosine phosphatase; phosphatase of activated
 C/Species: Homo sapiens (man)
 C/Date: 03-Nov-1995 #sequence_revision 02-Jul-1996 #text_change 11-Jun-1999
 C/Accession: A57126
 R/Rohan, P.J.; Davis, P.; Moskaluk, C.A.; Kearns, M.; Krutzech, H.; Stenblist, U.; Kelly
 Science 259, 1763-1766, 1993
 A/Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
 A/Reference number: A57126; MUID:93206122; PMID:7681221
 A/Accession: A57126
 A/Status: nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-314 <ROH>
 A/Cross-references: GB:111329; NID:G559539; PIDN:AAA50779.1; PID:G292376
 C/Genetics:
 A/Gene: GDB:DUSP2
 A/Cross-references: GDB:139200
 A/Map position: 2q11-2q11

A/Description: catalyzes the hydrolysis of peptidyl-phosphoserine, -phosphothreonine, and
 C/Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
 C/Keywords: nucleus, phosphoprotein, phosphoric monoester hydrolase
 F/140-160/Region: nuclear location signal
 F/180-311/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
 F/257/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/263/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.2%; Score 418; DB 1; Length 314;
 Best Local Similarity 36.0%; Pred. No. 2e-19;
 Matches 109; Conservative 45; Mismatches 105; Indels 44; Gaps 8;

RESULT 7
B57126 dual specificity phosphatase (EC 3.1.3.-) 2 - mouse
N:Alternate names: mitogen-induced nuclear protein-tyrosine-phosphatase PAC-1
C:Species: Mus musculus (house mouse)
C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #ext_change 24-Apr-1998
C:Accession: B57126
R:Kohan, P.U.; Davis, P.; Moshaluk, C.A.; Kearns, M.; Kruttsch, H.; Siebenlist, U.; Kel
Science 259, 1763-1766, 1993
A>Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
A:Reference number: A57126; MUID:93206122; PMID:7681221
A:Accession: B57126
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-314 <ROH>
A:Cross-references: Gb:L11330
C:Superfamily: dual specificity phosphoprotein
C:Keywords: nucleus, phosphoprotein, phosphoric monoester hydrolase
F:189-311/Domains: VHA-type dual specificity phosphoprotein phosphatase homology <VHA>
F:257/Active site: Cys (phosphocysteine intermediate) #status predicted
F:263/Binding site: substrate phosphate (Arg) #status predicted

A56947
dual specificity phosphatase (EC 3.1.3.-) HVM2 - rat
N:Alternate names: mitogen-activated protein kinase phosphatase 2
C:Species: Rattus norvegicus (Norway rat)
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-1999
C:Accession: A56947
R:Miller-Press, A.; Rim, C.S.; Yao, H.; Robertson, M.S.; Stork, P.J.S.
J. Biol. Chem. 270, 14587-14596, 1995
A:Title: A novel mitogen-activated protein kinase phosphatase. Structure, expression,
A:Reference number: A56947; MUID:95301550; PMID:7782322
A:Accession: A56947
A:Status: preliminary
A:Stature: preliminary
A:Molecule type: mRNA
A:Residues: 1-393 <MIS>
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity
C:Keywords: phosphoprotein; phosphoric monoester hydrolase
F:205-333/domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>
F:129/Active site: Cys (phosphocysteine intermediate) #status predicted
F:185/Binding site: substrate phosphate (Arg) #status predicted

```

Query Match      11.5%; Score 393.5; DB 2; Length 393;
Best Local Similarity 29.6%; Pred. No. 1e-17;
Matches 115; Conservative 64; Mismatches 153; Indels 57; Gaps 13;

QY      16 VALLSGTEKVEKLLIDSRPFVEYNTSHILEAININCSKLMKRRRLQODKYLITELLQHSAKH 75
      Db      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
      38 LEILLGG--KCLLLDRCPEFLHSAAGYINGVNVHCNTITVRR--AKGSVSLQILL--DAEE 92
      QY      76 KV-----DIDPSQKVVVVYDQSSQDVASLSLSDCLTYLLGKLEKSPNSVHLLAGFAEFS 129
      Db      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
      93 EVRPECALASTLSSTSTMTQPARRESPGQGSVAGRAALRRNAERT--DILCKGYERFS 150
      QY      130 RCFPELCEGKSTLVPTCISQCLPvani-----GPTRLPLMYLGGCQR 172
      Db      151 SEYEPFCSKTRAL--AAIPPEVPPTNESLDLGSSCGTFLHOGGVELLPFLVIGSAV 208
      QY      173 DVLNKEELIQONGIGVYNASTCPKPDFIEBSHFLKVPVNDSCCEKILPWLDRKSDPIEK 232
      Db      209 HAARDMDLALGITALLINVSDDCPN--HFEGHYQYCKIPVEDNHKADISSWFMALEYIDA 267
      QY      233 AKASNGCVLVHCLAGISRSATTAIAIYIMKRMDSLDEARFVYKERKPTISPNFNLGOLL 292
      Db      268 VKDGRGVLVHCOAGISRSATITCLAYIMKKRVLEAEFVKQRRSIIISPFSTMGOLL 327
      QY      293 DYERKTIKNQTAGSGPSPSKKLHLKEKNPEVPAVSREGGCKSETPLSPPCDASATSEAAQ 352
      Db      328 QFESQVLTTSACAFAAS-----PQGP-----LREKGAIPPT-----TSQFVFS 366
      QY      353 RPVHPASVSVSVSVQPSLLEDSPLVQALS 381
      Db      367 FPIV-SVGVHAAPSNNLPYL--HSPITTSFS 392

RESULT 9
I38890
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 5 - human
N:Alternate names: dual specificity phosphatase B23; dual-specificity phosphatase hVH-3,
C:Species: Homo sapiens (man)
C>Date: 16-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 11-Jun-1999
A:Accession: I38890; A5513
R:Kwak, S.P.; Dixon, J.E.
J. Biol. Chem. 270, 1156-1160, 1995
A>Title: Multiple dual specificity protein tyrosine phosphatases are expressed and regul
A:Reference number: A55432; MUID:5138103; PMID:7836374
A:Accession: I38890
A:Molecule type: mRNA
A:Residues: 1-384 <RES>
A:Cross-references: EMBL:U16996; NID:9642012; PIDN:AA06261.1; PID:g642013
A:Experimental source: Placenta
R:Shibashhi, T.; Bottaro, D.P.; Michieli, P.; Kelley, C.A.; Aaronson, S.A.
J. Biol. Chem. 269, 29897-29902, 1994
A>Title: A novel dual specificity phosphatase induced by serum stimulation and heat shock
A:Reference number: A55113; MUID:950508459; PMID:7961985

```


C/Accession: T18915
 R/Lloyd, C.
 submitted to the EMBL Data Library, November 1996
 A/Reference number: Z19044
 A/Accession: T18915
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-272 <MTL>
 A/Cross-references: EMBL:Z81461; PIDN:CA803837.1; GSPDB:GN00019; CESP:C04F12.8
 A/Experimental source: clone C04F12
 C/Genetics:
 A/Accession: CESP:C04F12.8
 A/Map position: 1
 A/Introns: 53/1; 84/3; 204/2

Query Match 7.0%; Score 239.5; DB 2; Length 272;
 Best Local Similarity 30.2%; Pred. No. 4.2e-08;
 Matches 77; Conservative 45; Mismatches 110; Indels 23; Gaps 8;
 QY 156 NIG---PTRLPNLYLGGQDVANKELIQONGIGVYLNASTYCPKPDFIPESHFLRPVND 211
 DB 4 NNGVLGQMSLNDHLVYS-GAGVLKPKDKIKQKIMVATTEEPS-TYMQGVDTWKIRI 61
 QY 212 NDSECEKILPWLDSKVDPIEKAKNSGCVLVHCLAGISRSATTAIAYIMKMDMSIDEAY 271
 DB 62 EDHRYARLNEHFDVADKINNVKERGKTLVHCHAGVSRASIVMLYLVKHEHMTLRQAY 121
 QY 272 RPYEKRPPTISPNFNLGQLDYEEKIKNOTGASGPKSKLKLHLKEPVPVASEG- 330
 DB 122 HYVAAPPIIRPVNGFMQWVDEKRLRG-----TASVKWQVPECMPIPDYVADI 174
 QY 331 --OKSEPTLSPCADSTSEAGORPVHPASV--PSYQSPQSLSEBSPLVQALSGHLISA 387
 DB 175 RMQINREISRHLSPASQASAKQAFASATYRPSVSA--SSLATSSLRAYSPSLPA 232
 QY 388 DRLEDNKTKRSPSL 402
 DB 233 SSL---ALTTYSYL 243

RESULT 13
 T21380
 hypothetical protein F26A3.4 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T21380
 R/McMurray, A.
 submitted to the EMBL Data Library, August 1996
 A/Reference number: Z19415
 A/Accession: T21380
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-226 <MTL>
 A/Cross-references: EMBL:Z78419; PIDN:CA801700.1; GSPDB:GN00019; CESP:F26A3.4
 A/Experimental source: clone F26A3
 C/Genetics:
 A/Accession: CESP:F26A3.4
 A/Map position: 1
 A/Introns: 117/2; 150/3; 186/3

Query Match 6.8%; Score 233.5; DB 2; Length 226;
 Best Local Similarity 28.6%; Pred. No. 7.7e-08;
 Matches 64; Conservative 47; Mismatches 80; Indels 33; Gaps 6;
 QY 160 TRILPNLYLGGQDVANKELIQONGIGVYLNASTYCPKPDFIPESHFLRPVNDSCFKI 219
 DB 15 SEIVPGFLI-CGVASLSEDEKHKIKTHINATTEVNLRLSGDIQRTKLMEIDTPQYI 73
 QY 220 LPMWDSKVDPIEKAKNSGCVLVHCLAGISRSATTAIAYIMKMDMSIDEAYRVEKRP 279
 DB 74 YPHLELSDQIGLALHAGKVLVHCVAGVSRASISLALFLKRCNNLEAHLKSKS 133
 QY 280 TISNENFLGQLDYEEKIKNOTGASGPKSKLKLHLKEPVPVASEGQKSETPLS 338

DB 134 WVRPNLGFWRQLIAVEQVKNKAS-----VRLVREADPEQLP----- 173
 QY 339 PPCADATSEAAAGRPVHPASVPSVQSPLEDSPLVALSG 382
 DB 174 ----DVYINIAI---PARPAS---PRDDPMKTPDEFRERRNSG 206

RESULT 14
 T16056
 hypothetical protein F13D11.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001
 C/Accession: T16056
 R/Fulton, L.
 submitted to the EMBL Data Library, November 1995
 A/Description: The sequence of C. elegans cosmid F13D11.
 A/Reference number: S69020
 A/Accession: T16056
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-186 <FUL>
 A/Cross-references: EMBL:U40939; NID:g1072175; PID:g1072176; PIDN:AAA81700.1; CESP:F13D11
 C/Genetics:
 A/Accession: CESP:F13D11.3
 A/Introns: 30/3; 57/3; 85/3; 125/3; 172/3

Query Match 6.6%; Score 225.5; DB 2; Length 186;
 Best Local Similarity 32.9%; Pred. No. 1.9e-07;
 Matches 57; Conservative 34; Mismatches 59; Indels 23; Gaps 5;
 QY 160 TRILPNLYL---GQGDVANKELIQONGIGVYLNASTYCPKPDFIPESHFLRPVNDSCFC 216
 DB 12 TVRPHLFLAGYGC-----ITPSLKQYNITHGVDTNKTXP--IKGLDRLEVPVDVTL 65
 QY 217 EKLPMWDSKVDPIEKAKNSGCVLVHCLAGISRSATTAIAYIMKMDMSIDEAYRVEKRP 276
 DB 66 AKITQYFEFVIVYIDAKQGHNTVYICAGVSSATITLYLWNTLSLEFAYLVQVQ 125
 QY 277 KRPTISPNFNLGQLDYEEKIKNOTGA---SGPKSKLKLHLKEPVPVAV 326
 DB 126 VRFIISPNIGFWRMITDEKQNGNASVELISGRMAR-----PVPSTV 167

RESULT 15
 T39698
 protein tyrosine phosphatase - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C/Accession: T39698
 R/Mood, V.; Skelton, J.; Churche, C.M.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, July 1999
 A/Reference number: Z21870
 A/Accession: T39698
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-330 <MOO>
 A/Cross-references: EMBL:AL109652; PIDN:CA851765.1; GSPDB:GN00067
 A/Experimental source: strain 972h-; cosmid c17A3
 C/Genetics:
 A/Accession: CESP:p1041
 A/Map position: 2

Query Match 6.5%; Score 223.5; DB 2; Length 330;
 Best Local Similarity 39.3%; Pred. No. 5.9e-07;
 Matches 53; Conservative 18; Mismatches 63; Indels 1; Gaps 1;
 QY 160 TRILPNLYLGGQDVANKELIQONGIGVYLNASTYCPKPDFIPESHFLRPVNDSCFKI 219
 DB 48 SEISKULYISSKWTABELVSTSKGIDYTLASMSINPNLS-VPEQHLMLQTEDESSQNI 106
 QY 220 LPMWDSKVDPIEKAKNSGCVLVHCLAGISRSATTAIAYIMKMDMSIDEAYRVEKRP 279

Db	107	LQYFEKSNKEIAPALSKNAKVLVHCFAGISRSVTLVAAVLMKENNMTTEALSHINERRS	166
Qy	280	TISPENFPLGOLDY	294
Db	167	GISPNANPLRLQRLRY	181

Search completed: June 21, 2004, 13:23:21
Job time : 23.6939 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:15:22 ; Search time 52.2647 Seconds
(without alignments)
4014.554 Million cell updates/sec

Title: US-10-029-345A-109
Perfect score: 3418
Sequence: 1 MAHEMIGCTIVTERLVALLE.....LGVKVSQSSFSQSGMEIIEVS 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.25.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3399	99.4	665	4	Q96N49	Q96N49 homo sapien
2	3379.5	98.9	662	4	Q81VT8	Q81VT8 homo sapien
3	3163	92.5	616	4	Q8N5T1	Q8N5T1 homo sapien
4	3079.5	90.1	660	11	Q920R2	Q920R2 mus musculus
5	2721.5	79.6	677	11	Q99MG6	Q99MG6 mus musculus
6	2380	69.6	622	11	Q99MG5	Q99MG5 mus musculus
7	1379.5	40.4	355	11	Q9AG16	Q9AG16 mus musculus
8	1327	38.8	625	4	Q86S88	Q86S88 homo sapien
9	1304	38.2	665	11	Q7TS29	Q7TS29 mus musculus
10	1038	30.4	300	11	Q9AG15	Q9AG15 mus musculus
11	891	26.1	206	11	Q8B2E4	Q8B2E4 mus musculus
12	611	17.9	657	5	Q8ST19	Q8ST19 caenorhabdi
13	610	17.8	143	4	Q96Q82	Q96Q82 homo sapien
14	497.5	14.6	606	5	Q8ST18	Q8ST18 caenorhabdi
15	487.5	14.3	367	13	Q7SZF3	Q7SZF3 brachydanto
16	469	13.7	382	13	Q7J2L9	Q7J2L9 gallus gall

17	469	13.7	483	11	Q8R3L3	Q8R3L3 mus musculus
18	458.5	13.4	439	5	Q8IG35	Q8IG35 caenorhabdi
19	449	13.1	383	13	Q8UW48	Q8UW48 fuigu rubrip
20	445	13.0	368	4	Q8NFT0	Q8NFT0 homo sapien
21	445	13.0	411	4	Q13649	Q13649 homo sapien
22	439.5	12.9	398	11	Q8BFT3	Q8BFT3 mus musculus
23	437	12.8	382	13	Q7J2L8	Q7J2L8 brachydanto
24	433.5	12.7	369	13	Q91790	Q91790 xenopus lae
25	431.5	12.6	369	13	Q90W58	Q90W58 xenopus lae
26	431	12.6	362	13	Q803B2	Q803B2 brachydanto
27	427.5	12.5	378	13	Q916B3	Q916B3 xenopus lae
28	424	12.4	367	11	Q64193	Q64193 rattus norv
29	421	12.3	367	11	Q63683	Q63683 rattus norv
30	415.5	12.2	353	13	Q42253	Q42253 gallus gall
31	415	12.1	360	13	Q7ZVL8	Q7ZVL8 brachydanto
32	413	12.1	318	11	Q80ZNI	Q80ZNI mus musculus
33	412	12.1	436	11	Q99KC2	Q99KC2 mus musculus
34	406	11.9	452	11	Q8K1S9	Q8K1S9 mus musculus
35	406	11.9	452	11	Q7TNL7	Q7TNL7 mus musculus
36	391.5	11.5	354	4	Q8N4A4	Q8N4A4 homo sapien
37	351.5	10.3	411	5	Q9SSV1	Q9SSV1 drosophila
38	351.5	10.3	411	5	Q9VVM5	Q9VVM5 drosophila
39	339	9.9	303	4	Q9NSW1	Q9NSW1 homo sapien
40	324.5	9.5	476	5	Q46122	Q46122 drosophila
41	319.5	9.3	476	5	Q9VHV8	Q9VHV8 drosophila
42	306	9.0	177	11	Q9CSL5	Q9CSL5 mus musculus
43	292.5	8.6	369	5	Q44128	Q44128 caenorhabdi
44	287	8.4	1045	5	Q9NKT1	Q9NKT1 drosophila
45	287	8.4	1193	5	Q8IMU8	Q8IMU8 drosophila

ALIGNMENTS

RESULT 1
Q96N49 PRELIMINARY; PRT; 665 AA.
ID Q96N49
AC Q96N49; 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ1411.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Ohshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.,
RA "NEO human cDNA sequencing project."
RT Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
RL EMBL; AK055973; BAB71060.1; -
DR HSSP; Q16828; IMR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR00340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; MAPK_phosph.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSRG; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE_3; 1.

DR PROSITE; PS00383; TYR PHOSPHATASE 1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hypothetical protein; Hydrolase.
 SQ SEQUENCE 665 AA; 73058 MW; 1EAEDDF08460DDF CRC64;

Query Match 99.4%; Score 3399; DB 4; Length 665;
 Best Local Similarity 99.5%; Pred. No. 3.1e-252;
 Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMIGTOIVTERLVALLESGETEKLIDSRPFVEYNTSHLEAININCKMKRRLOQ 60
 DB 1 MAHEMIGTOIVTERLVALLESGETEKLIDSRPFVEYNTSHLEAININCKMKRRLOQ 60
 QY 61 DKVLITELLQHSKAKKVIDDCQKVVYVDQSSQVVASLSDCFVLVLGLKLEKSFNSVHL 120
 DB 61 DKVLITELLQHSKAKKVIDDCQKVVYVDQSSQVVASLSDCFVLVLGLKLEKSFNSVHL 120
 QY 121 LAGGFAPFRCPPGLCEGKSTLVPTCISQPCLPVANIGPTILPNLYLGGORDVLANKEIM 180
 DB 121 LAGGFAPFRCPPGLCEGKSTLVPTCISQPCLPVANIGPTILPNLYLGGORDVLANKEIM 180
 QY 181 QONGIGYVLNASTYTPKPDFIPESHFLRPVNDSEFCETILPMLDKSVDFTEKAKANGCV 240
 DB 181 QONGIGYVLNASTYTPKPDFIPESHFLRPVNDSEFCETILPMLDKSVDFTEKAKANGCV 240
 QY 241 LVHCLAGISRATTAIAIYIMKMDMSLDEAYRFYKERRPTISPNFNLGQLDYEKKIKN 300
 DB 241 LVHCLAGISRATTAIAIYIMKMDMSLDEAYRFYKERRPTISPNFNLGQLDYEKKIKN 300
 QY 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGQKSETPLSPCADSATSEAGRPVHPASV 360
 DB 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGQKSETPLSPCADSATSEAGRPVHPASV 360
 QY 361 PSVSVQPSLLEDSPLVQALSGHLTSLADRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
 DB 361 PSVSVQPSLLEDSPLVQALSGHLTSLADRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
 QY 421 SSSSDALEYKPSSTITLDGNTKLCOFSPVOELSEQTPETSPDKEASIPKQLQTARPSDQ 480
 DB 421 SSSSDALEYKPSSTITLDGNTKLCOFSPVOELSEQTPETSPDKEASIPKQLQTARPSDQ 480
 QY 481 SKRLHSYRTSSSGTAQRSLISPLHRSQVSDNYHTSPLFGJLSTGOHLTKSAGLGLKGMH 540
 DB 481 SKRLHSYRTSSSGTAQRSLISPLHRSQVSDNYHTSPLFGJLSTGOHLTKSAGLGLKGMH 540
 QY 541 SDILAQSTSTPLTSSWYFATESSHFYASAIYGGASAYSQSLPTCGDYYSVRRR 600
 DB 541 SDILAQSTSTPLTSSWYFATESSHFYASAIYGGASAYSQSLPTCGDYYSVRRR 600
 QY 601 QKPSRADRRKRWHESSPEKQFKRRSCOMFEGSINSNRSRELKGVSGSSFSGSM 660
 DB 601 QKPSRADRRKRWHESSPEKQFKRRSCOMFEGSINSNRSRELKGVSGSSFSGSM 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 2

Q81VT8 PRELIMINARY; PRT; 662 AA.

AC Q81VT8;
 DT 01-MAR-2003 (Tremblrel_23, Created)
 DT 01-MAR-2003 (Tremblrel_23, Last sequence update)
 DE 01-OCT-2003 (Tremblrel_23, Last annotation update)
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eultheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Duodenum;
 RA Strausberg R.;
 RL Submitted (DEC-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; BC042101; AAH42101.1; --
 DR GO; GO:0017017; P-MAP kinase phosphatase activity; IEA.
 DR GO; GO:0006470; P-protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000383; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; MARKPHATASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS0206; RHODANES 3; 1.
 DR PROSITE; PS00383; TYR PHOSPHATASE 1; 1.
 DR PROSITE; PS50056; TYR PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR PHOSPHATASE_DUAL; 1.
 SQ SEQUENCE 662 AA; 72818 MW; 350534EF0652B98F CRC64;

Query Match 98.3%; Score 3379.5; DB 4; Length 662;
 Best Local Similarity 99.2%; Pred. No. 9.6e-251;
 Matches 660; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 MAHEMIGTOIVTERLVALLESGETEKLIDSRPFVEYNTSHLEAININCKMKRRLOQ 60
 DB 1 MAHEMIGTOIVTERLVALLESGETEKLIDSRPFVEYNTSHLEAININCKMKRRLOQ 60
 QY 61 DKVLITELLQHSKAKKVIDDCQKVVYVDQSSQVVASLSDCFVLVLGLKLEKSFNSVHL 120
 DB 61 DKVLITELLQHSKAKKVIDDCQKVVYVDQSSQVVASLSDCFVLVLGLKLEKSFNSVHL 120
 QY 121 LAGGFAPFRCPPGLCEGKSTLVPTCISQPCLPVANIGPTILPNLYLGGORDVLANKEIM 180
 DB 121 LAGGFAPFRCPPGLCEGKSTLVPTCISQPCLPVANIGPTILPNLYLGGORDVLANKEIM 180
 QY 181 QONGIGYVLNASTYTPKPDFIPESHFLRPVNDSEFCETILPMLDKSVDFTEKAKANGCV 240
 DB 181 QONGIGYVLNASTYTPKPDFIPESHFLRPVNDSEFCETILPMLDKSVDFTEKAKANGCV 240
 QY 241 LVHCLAGISRATTAIAIYIMKMDMSLDEAYRFYKERRPTISPNFNLGQLDYEKKIKN 300
 DB 241 LVHCLAGISRATTAIAIYIMKMDMSLDEAYRFYKERRPTISPNFNLGQLDYEKKIKN 300
 QY 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGQKSETPLSPCADSATSEAGRPVHPASV 360
 DB 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGQKSETPLSPCADSATSEAGRPVHPASV 360
 QY 361 PSVSVQPSLLEDSPLVQALSGHLTSLADRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
 DB 361 PSVSVQPSLLEDSPLVQALSGHLTSLADRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
 QY 421 SSSSDALEYKPSSTITLDGNTKLCOFSPVOELSEQTPETSPDKEASIPKQLQTARPSDQ 480
 DB 421 SSSSDALEYKPSSTITLDGNTKLCOFSPVOELSEQTPETSPDKEASIPKQLQTARPSDQ 480
 QY 481 SKRLHSYRTSSSGTAQRSLISPLHRSQVSDNYHTSPLFGJLSTGOHLTKSAGLGLKGMH 540
 DB 481 SKRLHSYRTSSSGTAQRSLISPLHRSQVSDNYHTSPLFGJLSTGOHLTKSAGLGLKGMH 540
 QY 541 SDILAQSTSTPLTSSWYFATESSHFYASAIYGGASAYSQSLPTCGDYYSVRRR 600
 DB 541 SDILAQSTSTPLTSSWYFATESSHFYASAIYGGASAYSQSLPTCGDYYSVRRR 600
 QY 601 QKPSRADRRKRWHESSPEKQFKRRSCOMFEGSINSNRSRELKGVSGSSFSGSM 660
 DB 601 QKPSRADRRKRWHESSPEKQFKRRSCOMFEGSINSNRSRELKGVSGSSFSGSM 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 3
ID Q8N5T1 PRELIMINARY; PRT; 616 AA.
AC Q8N5T1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Lung;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC031643; AAH31643.1; -
DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro: IPR000343; MAPK_phosph.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 616 AA; 67636 MW; 2CBOB14482F2AD72 CRC64;
Query Match 92.5%; Score 3163; DB 4; Length 616;
Best Local Similarity 99.5%; Pred. No. 3.7e-234;
Matches 613; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 50 CSKLMKRLQODQVLTTELQHSKAKKVIDCSQKVVVYQSSQDVASLSDDCLTYLLG 109
1 CSKLMKRLQODQVLTTELQHSKAKKVIDCSQKVVVYQSSQDVASLSDDCLTYLLG 60
QY 110 KLEKSFNSVHLLGGFAEFSCFPGLCCEKSTLVPTCISQPCLVANIGFTRILPNLYLG 169
61 KLEKSFNSVHLLGGFAEFSCFPGLCCEKSTLVPTCISQPCLVANIGFTRILPNLYLG 120
QY 170 QCRDVLNKLQONGIGYVNAASYTCRPFIPESHFLRPVNDSPCEKILPWLDSVDF 229
121 QCRDVLNKLQONGIGYVNAASYTCRPFIPESHFLRPVNDSPCEKILPWLDSVDF 180
QY 230 IERAKASNGCVLVHCLAGISRATIAIYIMKMDMSLDENRPFVKEKRPITISPNENFLG 289
181 IERAKASNGCVLVHCLAGISRATIAIYIMKMDMSLDENRPFVKEKRPITISPNENFLG 240
QY 290 QILDYEKKINQNGASGPKSKLKLHLEKNEPVPAVSEGGSEMPPLSPCCDSATSEA 349
241 QILDYEKKINQNGASGPKSKLKLHLEKNEPVPAVSEGGSEMPPLSPCCDSATSEA 300
QY 350 AGORPVPAVSPVSPVQPSLLEDSPLVQALSGHLHLSADRLSDSNKLKRSFSLDIKSVY 409
301 AGORPVPAVSPVSPVQPSLLEDSPLVQALSGHLHLSADRLSDSNKLKRSFSLDIKSVY 360
QY 410 SASMAASLHGSSSEDALEYYKQSTTLDGTNKLQFSPVQELSEQIPETISPDYKESAIK 469
361 SASMAASLHGSSSEDALEYYKQSTTLDGTNKLQFSPVQELSEQIPETISPDYKESAIK 420
QY 470 KLTQARSPDSOSKRLHVSRTSSGTAQRSLSPLRHSGSVEDNPHYSFLGLSTSOQHLT 529
421 KLTQARSPDSOSKRLHVSRTSSGTAQRSLSPLRHSGSVEDNPHYSFLGLSTSOQHLT 480
QY 530 KSAAGLGLKGMHSDILAPQTSPTSLTSSWYPATBESHFYSAIYGGASAYSCSOLPT 589
481 KSAAGLGLKGMHSDILAPQTSPTSLTSSWYPATBESHFYSAIYGGASAYSCSOLPT 540

QY 590 CGDQVYVRRROKPSRADRSRWSHESPKEQFKRRSCOMERGESIMSENRREELGKY 649
DB 541 CGDQVYVRRROKPSRADRSRWSHESPKEQFKRRSCOMERGESIMSENRREELGKY 600
QY 650 GSQSFSGSMELTEIVS 665
DB 601 GSQSFSGSMELTEIVS 616
RESULT 4
ID Q920R2 PRELIMINARY; PRT; 660 AA.
AC Q920R2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MAP kinase phosphatase-7.
GN DUSP16 OR 3830417M17RIK OR MKP-7.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
RT "MKP-7, a Novel Mitogen-activated Protein Kinase Phosphatase,
RT Functions as a Shuttle Protein."
RL J. Biol. Chem. 276:39002-39011(2001).
DR EMBL: AB052157; BAB47240.1; -
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1917936; Dusp16.
DR GO: GO:0005737; C:cytoplasm; IDA.
DR GO: GO:0005634; C:nucleus; IDA.
DR GO: GO:0005515; P:protein binding; IPI.
DR GO: GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ SEQUENCE 660 AA; 72695 MW; DB609FCDAD4AA309 CRC64;
Query Match 90.1%; Score 3079.5; DB 11; Length 660;
Best Local Similarity 90.1%; Pred. No. 1.1e-227;
Matches 599; Conservative 28; Mismatches 33; Indels 5; Gaps 3;
QY 1 MAHEMIGTOIVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
1 MAHEMIGTOIVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
QY 61 DKVLITELQHSKAKKVIDCSQKVVVYQSSQDVASLSDDCLTYLLGLKLEKSFNSVH 120
61 DKVLITELQHSKAKKVIDCSQKVVVYQSSQDVASLSDDCLTYLLGLKLEKSFNSVH 120
QY 121 LAGFAEFSCFPGLCCEKSTLVPTCISQPCLVANIGFTRILPNLYLGQCRDVLNKL 180
121 LAGFAEFSCFPGLCCEKSTLVPTCISQPCLVANIGFTRILPNLYLGQCRDVLNKL 180
QY 121 LAGFAEFSCFPGLCCEKSTLVPTCISQPCLVANIGFTRILPNLYLGQCRDVLNKL 180
121 LAGFAEFSCFPGLCCEKSTLVPTCISQPCLVANIGFTRILPNLYLGQCRDVLNKL 180
QY 181 QONGIGYVNAASYTCRPFIPESHFLRPVNDSPCEKILPWLDSVDFIERAKASNGCV 240
181 QONGIGYVNAASYTCRPFIPESHFLRPVNDSPCEKILPWLDSVDFIERAKASNGCV 240
DB 181 QONGIGYVNAASYTCRPFIPESHFLRPVNDSPCEKILPWLDSVDFIERAKASNGCV 240

```

QY 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTTISPNFNFGQLLDYEKTKIN 300
Db 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTTISPNFNFGQLLDYEKTKIN 300
QY 301 QTASGPKSKLKLHLHLEKNEPVPAVSEGGQSETPSPCCADSATSEAGORPVHPASV 360
Db 301 QTASGPKSKLKLHLHLEKNEPVPAVSEGGQSETPSPCCADSATSEAGORPVHPASV 360
QY 361 PSVPSVQPSLLEDSPLVQALSGHLASDLLEDSNKLKSPSLDIKVSYSASMAASLHGF 420
Db 361 PSVPSVQPSLLEDSPLVQALSGHLASDLLEDSNKLKSPSLDIKVSYSASMAASLHGF 420
QY 421 SSEDALERYKPTSLDGTNKLQCPSPVOELSEQTPETSPDKEASIPKQLQTARPDSQ 480
Db 421 SSEDALERYKPTSLDGTNKLQCPSPVOELSEQTPETSPDKEASIPKQLQTARPDSQ 480
QY 481 SKRLHSVRTSSGTAORSLSPLRSGSVEENHTSFLPGLSTSQOHLTKSAGLGLKGMH 540
Db 481 SKRLHSVRTSSGTAORSLSPLRSGSVEENHTSFLPGLSTSQOHLTKSAGLGLKGMH 540
QY 541 SDIILAPQSTSTSLTSSWTFATESSHFYASAIYGSASYSAYSCSOLPTCGDQVSVRR 600
Db 541 SDIILAPQSTSTSLTSSWTFATESSHFYASAIYGSASYSAYSCSOLPTCGDQVSVRR 600
QY 601 QKPSDRADSRSMHEESPEKQFRRSCOMEFGESIMSENRREELGKYSGSFGSGME 660
Db 601 QKPSDRADSRSMHEESPEKQFRRSCOMEFGESIMSENRREELGKYSGSFGSGME 660
QY 661 IIEVS 665
Db 661 IIEVS 660

```

RESULT 5

```

QY 099MG6 PRELIMINARY; PRT; 677 AA.
Db 099MG6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Map kinase phosphorylation update)
DE Map kinase phosphorylation-M A1 isoform.
OS DUSP16 OR 3830417M17RIK.
CS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=BALB/c;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages."
RT Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL; AF345951; AAK35052.1; -.
DR HSSP; 016828; IMKP.
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005634; Nucleus; IDA.
DR GO; GO:0005515; Protein binding; IPI.
DR GO; GO:000188; Phosphatase; IDA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR000343; MAPK phosphatase.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR003871; TYR phosphatase.
DR Pfam; PF00782; DSCP; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHTRASE.
DR SMART; SM00195; DSCP; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.

```

```

DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ SEQUENCE 677 AA; 74550 MW; 8B6D5B7096C8C2FC CRC64;

```

```

Query Match 79.6%; Score 2721.5; DB 11; Length 677;
Best Local Similarity 84.4%; Pred. No. 3.5e-200;
Matches 531; Conservative 36; Mismatches 57; Indels 5; Gaps 3;

```

```

QY 1 MAHEMIGQIVTERIVALLESSTKVLIDSRPFVEYNTSHLEAININCKMKRLQ 60
Db 1 MAHEMIGQIVTERIVALLESSTKVLIDSRPFVEYNTSHLEAININCKMKRLQ 60
QY 61 DKVITTELQHSAGHKVDICQSVVVYDDSSQDVASLSDCHLYLGLKESFNSVHL 120
Db 61 DKVITTELQHSAGHKVDICQSVVVYDDSSQDVASLSDCHLYLGLKESFNSVHL 120
QY 121 LAGGPAEFSRCFPCGCKSTLPTCISQPCLPVANIIGTRILPMLYLCCQDVLNKL 180
Db 121 LAGGPAEFSRCFPCGCKSTLPTCISQPCLPVANIIGTRILPMLYLCCQDVLNKL 180
QY 181 QONGIGVILNASTCPKDPDIPESHFLRVVNDSCFKILPWLKSVDFIEKAKASNCV 240
Db 181 QONGIGVILNASTCPKDPDIPESHFLRVVNDSCFKILPWLKSVDFIEKAKASNCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTTISPNFNFGQLLDYEKTKIN 300
Db 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTTISPNFNFGQLLDYEKTKIN 300
QY 301 QTASGPKSKLKLHLHLEKNEPVPAVSEGGQSETPSPCCADSATSEAGORPVHPASV 360
Db 301 QTASGPKSKLKLHLHLEKNEPVPAVSEGGQSETPSPCCADSATSEAGORPVHPASV 360
QY 361 PSVPSVQPSLLEDSPLVQALSGHLASDLLEDSNKLKSPSLDIKVSYSASMAASLHGF 420
Db 361 PSVPSVQPSLLEDSPLVQALSGHLASDLLEDSNKLKSPSLDIKVSYSASMAASLHGF 420
QY 421 SSEDALERYKPTSLDGTNKLQCPSPVOELSEQTPETSPDKEASIPKQLQTARPDSQ 480
Db 421 SSEDALERYKPTSLDGTNKLQCPSPVOELSEQTPETSPDKEASIPKQLQTARPDSQ 480
QY 481 SKRLHSVRTSSGTAORSLSPLRSGSVEENHTSFLPGLSTSQOHLTKSAGLGLKGMH 540
Db 481 SKRLHSVRTSSGTAORSLSPLRSGSVEENHTSFLPGLSTSQOHLTKSAGLGLKGMH 540
QY 541 SDIILAPQSTSTSLTSSWTFATESSHFYASAIYGSASYSAYSCSOLPTCGDQVSVRR 600
Db 541 SDIILAPQSTSTSLTSSWTFATESSHFYASAIYGSASYSAYSCSOLPTCGDQVSVRR 600
QY 601 QKPSDRADSRSMHEESPEKQFRRSCOMEFGESIMSENRREELGKYSGSFGSGME 660
Db 601 QKPSDRADSRSMHEESPEKQFRRSCOMEFGESIMSENRREELGKYSGSFGSGME 660
QY 661 IIEVS 665
Db 661 IIEVS 660

```

RESULT 6

```

QY 099MG6 PRELIMINARY; PRT; 622 AA.
Db 099MG6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Map kinase phosphorylation update)
DE Map kinase phosphorylation-M A2 isoform.
OS DUSP16 OR 3830417M17RIK.
CS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=BALB/c;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK

```

RT Activation in Macrophages.
 RL Mol. Cell. Biol. 20:6999-7009 (2001).
 DR EMBL: AF345952; AAK35053.1; -.
 DR HSSP: Q16828; IMKP.
 DR MGD: MGI:1917936; Dusp16.
 DR GO: GO:0005737; C:cytoplasm; IDA.
 DR GO: GO:0005634; C:nucleus; IDA.
 DR GO: GO:0005515; F:protein binding; IPI.
 DR GO: GO:000188; P:inactivation of MAPK; IDA.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSpC; 1.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00450; DSpC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS50206; RHODANSE 3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE 1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE 2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR HydroLase: Kinase.
 KW HydroLase: Kinase.
 SQ SEQUENCE 622 AA; 68672 MW; 6C0CA4E6E909B98 CRC64;

Query Match 69.6%; Score 2380; DB 11; Length 622;
 Best Local Similarity 75.7%; Pred. No. 5.2e-174;
 Matches 476; Conservative 36; Mismatches 57; Indels 60; Gaps 4;

QY 1 MAHEMIGTQVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 DB 1 MAHEMIGTQVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 QY 61 DKVLITELIGHSAKHKVIDCSQKVVVYDSSQDVASLSDDCFYLVLGLEKSFNSVHL 120
 DB 61 DKVLITELIGHSAKHKVIDCSQKVVVYDSSQDVASLSDDCFYLVLGLEKSFNSVHL 120
 QY 121 LAGGFARFCRPGLCGKSTLVPTCISQPCLVANIGPRLIPNLVYLGQDRVNLKELI 180
 DB 121 LAGGFARFCRPGLCGKSTLVPTCISQPCLVANIGPRLIPNLVYLGQDRVNLKELI 180
 QY 121 LA-----DLM 125
 DB 121 LA-----DLM 125
 QY 181 QONGIGYVLAASYCPKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 DB 181 QONGIGYVLAASYCPKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 QY 126 QONGIGYVLANASTCEKDPDIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 185
 DB 126 QONGIGYVLANASTCEKDPDIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 185
 QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLDYEKKIKN 300
 DB 186 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLDYEKKIKN 245
 QY 301 QTAASGPKSLKLIHLEKNEPVPVNSGQKSETPLSPCADSATSEAGQRPVPVASY 360
 DB 246 QTAASGPKSLKLIHLEKNEPVPVNSGQKSETPLSPCADSATSEAGQRPVPVASY 302
 QY 361 PSVPSPQSLLEDDSPVQALSGHLASDLRLEDSNKLKRSFLDIKSYVASMAASLHGF 420
 DB 303 -SVPSLQPSLLEDDSPVQALSGHLASDLRLEDSNKLKRSFLDIKSYVASMAASLHGF 361
 QY 421 SSSEDALEYYKPSSTLIDGNTNKLQFSPVQSLSEQTPETSPDKKEASI PKKLQTAAPSDQ 480
 DB 362 -SSEDALEYYKPSSTLIDGNTNKLQFSPVQSLSEQTPETSPDKKEASI PKKLQTAAPSDQ 420
 QY 481 SKRLHVSRRSSGTAQRSLSPHRSQSVEDNHTVSTLPGSLSTQOHLTASAGLGLKGM 540
 DB 421 VTLHSHVRGSSGTAQRSLSPHRSQSVEDNHTVSTLPGSLSTQOHLTASAGLGLKGM 480
 QY 541 SDLIAGQTSPTSLTSSWYFATSSSHFYASAGIYGSASAYSCSOLPTCGDOVYVRRR 600
 DB 481 SDLIAGQTSPTSLTSSWYFATSSSHFYASAGIYGSASAYSCSOLPTCGDOVYVRRR 540
 QY 601 QKPSDRAISRSHWESPFQKFRSSCQ 629
 DB 541 QKPSDRAISRSHWESPFQKFRSSCQ 629

RESULT 7
 ID 09AG16 PRELIMINARY; PRT; 355 AA.
 AC 09AG16;
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Map Kinase phosphatase-M B1 isoform.
 GN DUSP16 OR 383041TM17RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RA Matsuguchi T., Musikacharen T., Johnson T.R., Kraft A.S.,
 RA Yoshikai Y.;
 RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
 RT Activation in Macrophages."
 RL Mol. Cell. Biol. 20:6999-7009 (2001).
 DR EMBL: AF345953; AAK35054.1; -.
 DR HSSP: Q16828; IMKP.
 DR MGD: MGI:1917936; Dusp16.
 DR GO: GO:0005737; C:cytoplasm; IDA.
 DR GO: GO:0005634; C:nucleus; IDA.
 DR GO: GO:0005515; F:protein binding; IPI.
 DR GO: GO:000188; P:inactivation of MAPK; IDA.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSpC; 1.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00450; DSpC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS50206; RHODANSE 3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE 1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE 2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW HydroLase: Kinase.
 SQ SEQUENCE 355 AA; 39502 MW; D52C29AE215CA285 CRC64;

Query Match 40.4%; Score 1379.5; DB 11; Length 355;
 Best Local Similarity 83.2%; Pred. No. 1.5e-97;
 Matches 273; Conservative 13; Mismatches 23; Indels 19; Gaps 2;

QY 1 MAHEMIGTQVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 DB 1 MAHEMIGTQVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 QY 61 DKVLITELIGHSAKHKVIDCSQKVVVYDSSQDVASLSDDCFYLVLGLEKSFNSVHL 120
 DB 61 DKVLITELIGHSAKHKVIDCSQKVVVYDSSQDVASLSDDCFYLVLGLEKSFNSVHL 120
 QY 121 LAGGFARFCRPGLCGKSTLVPTCISQPCLVANIGPRLIPNLVYLGQDRVNLKELI 180
 DB 121 LAGGFARFCRPGLCGKSTLVPTCISQPCLVANIGPRLIPNLVYLGQDRVNLKELI 180
 QY 181 QONGIGYVLAASYCPKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 DB 181 QONGIGYVLAASYCPKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKE-----KRPITSPNPNFLG 289
 DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKE-----KRPITSPNPNFLG 289
 QY 290 QLLDYEKKIKNTG-----ASGPKS 309
 DB 301 AKMNLERALCRRTGPRSWARWASPAS 328

RESULT 8

Q86SS8

ID Q86SS8 PRELIMINARY; PRT; 625 AA.

AC Q86SS8

DT 01-JUN-2003 (Tremblrel. 24, Created)

DT 01-UN-2003 (Tremblrel. 24, Last sequence update)

DT 01-0CT-2003 (Tremblrel. 25, Last annotation update)

DE Similar to dual specificity phosphatase 8.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain, and Astrocytoma;

RA Strausberg R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR GO: GO:0006470; F:MAP kinase phosphatase activity; IEA.

DR GO: GO:0017017; P:Protein amino acid dephosphorylation; IEA.

DR InterPro: IPR000340; DS_phosphatase.

DR InterPro: IPR008343; MAPK_phosph.

DR InterPro: IPR002965; P-rich extensn.

DR InterPro: IPR001763; Rhodanese-like.

DR Pfam: PF00782; DSPC, 1.

DR PRINTS: PR01764; MAPKPHPTASE.

DR PRINTS: PR01217; PRICHEXTENS.

DR SMART: SM00450; RHOD; 1.

DR PROSITE: PS00206; RHODANES 3; 1.

DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.

DR PROSITE: PS00564; TYR_PHOSPHATASE_2; 1.

DR PROSITE: PS00564; TYR_PHOSPHATASE_DUAL; 1.

SQ SEQUENCE 625 AA; 65826 MW; C7C808407B724FFC CRC64;

Query Match 38.8%; Score 1327; DB 4; Length 625;

Best Local Similarity 45.1%; Pred. No. 4e-93;

Matches 314; Conservative 94; Mismatches 185; Indels 104; Gaps 21;

1 MAHEMIGTOIV-TERIVALLSGETEKLILDSRPFEVNTSHLEAININCSKLMKRLQ 59
1 MAGDRILPRKYMADKSLASLRGPGGPIVLDISRFVENSVMHVLASVNICCSKLVKRLQ 60
60 QDKVITLIELIQHSAKHKVDIDCSQKVVVYDSSQVAVSLSDCFILVLLGLKLEKSFNSVH 119
61 QCKVTIABLIOPAAISOVEATEPQVVVVYDSTRDASVLAADSLFSLILSKLDCGCFDPSVA 120
120 ILAAGPAEFSRCFELCEGK-STLVPTCISQPCLPVANIGPTRLIPNLVYGGCORVLANKE 178
121 ILTGFATPSSCFELCEGKPAALLPMSLSQPCLPVPSVGLTRILPHLYLGSQKVDVANKD 180
179 LIOQNGIGVYNAASYTCCKPDPIPESHLARPVNDSCFEKLLPWLDKSVDFIEKAKASNG 238
181 LMTQNGISYVNAASCPKDPFICSRFMRVPINDNYCEKLLPWLDKSIFIDIKAKLSISC 240
239 CVLVHCLAGISRSATIALAYIMKEMDSIDEAYFVFEKERTTSPNNFNGOLLDVYEKTI 298
241 QVIYHCLAGISRSATIALAYIMKEMDSIDEAYFVDRPSPISPNNFNGOLLDVYEKTI 300
299 KNOVGASGPKSKLILHLEKPN---EPVPAVSEGGQKSETPPLSPCCADSNAT---SEAAGOR 353
301 KLLALALQDDPG-----TPSGTPBPSPPAAGAPLPRLP---PPTSESAATGNAALARG 350
354 PVHPSVPSVPSVPSVPSVPSVPSVPSVPSVPSVPSVPSVPSVPSVPSVPSVPSVPSV 413
351 GLASGGBPPAPPPPPA---TSAIQGLRGHLHSLSDRLQDNTNRKRKFSLDIKSA----- 401
414 AASLHGSSSEDALEYYKSTTLDTGN-----KLQCF-SP---VOELSEQTPETSP 461
402 -----YAPSRBPDPGPPDPGAPKCKLSDSPSGAALGLSSPSD-SFD 444

RESULT 9

Q7TS29

ID Q7TS29 PRELIMINARY; PRT; 665 AA.

AC Q7TS29

DT 01-0CT-2003 (Tremblrel. 25, Created)

DT 01-0CT-2003 (Tremblrel. 25, Last sequence update)

DT 01-0CT-2003 (Tremblrel. 25, Last annotation update)

DE Dusp8 protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Stapleton L., Maruska K., Farmer A.A., Rubin G.M., Hong L.,

RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Scheetz T.E.,

RA Raha S.S., Lonellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McMan P.J., McKernan K.J., Gale J.A., Gunaratne P.H.,

RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood U., Schmutz J., Myers R.M., Buterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.,

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903 (2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Strausberg R.;

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC052705; AAH52705.1; -

SQ SEQUENCE 665 AA; 69021 MW; 9166E36A8835249F CRC64;

Query Match 38.2%; Score 1304; DB 11; Length 665;

Best Local Similarity 43.5%; Pred. No. 2.5e-91;

Matches 319; Conservative 101; Mismatches 176; Indels 138; Gaps 26;

1 MAHEMIGTOIV-TERIVALLSGETEKLILDSRPFEVNTSHLEAININCSKLMKRLQ 59
1 MAGDRILPRKYMADKSLASLRGPGGPIVLDISRFVENSVMHVLASVNICCSKLVKRLQ 60
60 QDKVITLIELIQHSAKHKVDIDCSQKVVVYDSSQVAVSLSDCFILVLLGLKLEKSFNSVH 119
61 QCKVTIABLIOPAAISOVEATEPQVVVVYDSTRDASVLAADSLFSLILSKLDCGCFDPSVA 120

QY 120 LLAGFAEFCRPGCEGKSTLVPT-CISQPCLPVANIGTRILPMLYLCCORDVINKR 178
 DB 121 ILTGFGATFSCFPGCEGKAPLPSMSLSQCLPVPVGLTILPHLVLGSKDVLNMD 180
 QY 179 LIQONGIGYVLNASTYCPKPDFPESHFLRPVNDSECEKILWLDKSDVFIKAKASNG 238
 DB 181 LMTQNGISYVLNASTNCPKPDFICESFMRIPINDVNCETLFWLDKSLIEFIDKALSSC 240
 QY 239 CVLVHCLAGISRSATTAIAYIMKMDMSLDEAYFVKEKPTSPNPFGLDDEYK 298
 DB 241 QVIVHCLAGISRSATTAIAYIMKMDMSLDEAYFVKEKPTSPNPFGLDDEYK 300
 QY 299 KNOTG--ASGPKSKLKLHLEKNEPEVPVASEGOKSETPLSPPCADSAT--SEA- 350
 DB 301 KLIALAQDTGP-----HIGTP-EPLMGPAAGIPLRLP--PSTSESATSEANAR 350
 QY 351 -----GORPHAPVSVVQPSLLEDSPVLQALSGHLSADRLSDNKLKSPSLD 403
 DB 351 EGSPSAGDAPL-PSTAPATSAQ-----QGLRGHLSSDRLOQDTNRLKRSFSLD 399
 QY 404 IKSVYSASMAALHGSSEDALEYKSTLTLDGNTKLCOFSPVQELSTQTEPDK 463
 DB 400 IKS-AVAPSRPDPFGPPDGEAPKCLKDSFGGTLGL-----PSPSPDP 445
 QY 464 EASIPK--KLQTAPEPDSQSKR--LHSVRTSSSGTAORELSPLHRSSEVEDNHTSFL 518
 DB 446 D-SVPECRRPPRRPPASSPAASPAHGLGDLNAGDTAKQ--TPRH----- 467
 QY 519 FGLSTSQOHLTKSAGL-----GLKGMHSDILAPQTSPLTSSWYFATESH--FY 567
 DB 488 -GLSA-----LSAPGLPQPGQPAQPGWVPLDSP--GTPSPDGMCFSPREGAQPGAVP 539
 QY 568 SA-----SATYGG-----SASYAVSCQPLPCGD 552
 DB 540 SAGRVYAGAPGPNSSSGGGGGGGGGGGSSSSSSSSSSSSSSSSSSSSSSSSSS 599
 QY 593 QVYVVRROKPSRADSRSMHESPEKQFKRRSCOMERGESI MSNRR-BELGVGS 651
 DB 600 SSSDLNRR-----DRTGMPEEPADAOFKRRSCOMEREG-MEGRAGBELALGK 651
 QY 652 QSSFGSGMEIIEVS 665
 DB 652 QTSFGSGVEIIEVS 665

RESULT 10
 Q9AG15 PRELIMINARY; PRT; 300 AA.

AC 09AG15;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Map kinase phosphatase-M B2 isoform.
 GN DUSP16 OR 3830417M17R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BA1B/c; Muslacharoen T., Johnson T.R., Kraft A.S.,
 RA Matsuguchi T., Yoshioka Y.,
 RA "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
 RT Activation in Macrophages.";
 RT Mol. Cell. Biol. 20:6999-7009 (2001).
 RL EMBL; AF345554; AAK35055.1; -.
 DR HSP; Q16828; MKP.
 DR MGD; MG1:1917936; Dusp16.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0000188; P:inactivation of MAPK; IDA.

DR InterPro; IPR000340; D5_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANASE_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydroxylase; Kinase.
 SQ SEQUENCE 300 AA; 33624 MW; 4C61846ACDF0F456 CRC64;

Query Match 30.4%; Score 1038; DB 11; Length 300;
 Best Local Similarity 66.5%; Pred. No. 2e-71;
 Matches 218; Conservative 13; Mismatches 23; Indels 74; Gaps 3;

QY 1 MAHEMGTQIVTERLVALLSGTEKYLIDSRPFVYNTSHILEAININCSKLMKRLQ 60
 DB 1 MAHEMGTQIVTESVALLSGTEKYLIDSRPFVYNTSHILEAININCSKLMKRLQ 60
 QY 61 DKVLITELQHSKHKVDIDCSQKVVYVQSSODVASLSDDCFVLVLGLKGFNSVHL 120
 DB 61 DKVLITELQHSKHKVDIDCSQKVVYVQSSODVASLSDDCFVLVLGLKGFNSVHL 120
 QY 121 LLAGFAEFCRPGCEGKSTLVPTCISQPCLPVANIGTRILPMLYLCCORDVINKELI 180
 DB 121 LA-----DLN 125
 QY 181 QONGIGYVLNASTYCPKPDFPESHFLRPVNDSECEKILWLDKSDVFIKAKASNGCV 240
 DB 126 QONGIGYVLNASTYCPKPDFPESHFLRPVNDSECEKILWLDKSDVFIKAKASNGCV 185
 QY 241 LVHCLAGISRSATTAIAYIMKMDMSLDEAYFVKE-----KPTSPNPFGL 289
 DB 186 LVHCLAGISRSATTAIAYIMKMDMSLDEAYFVKE-----KPTSPNPFGL 245
 QY 290 QLDYEKKIKNOTG-----ASGPKS 309
 DB 246 AKWNLERALCRRGPGRSWARMASPAS 273

RESULT 11
 Q8BZE4 PRELIMINARY; PRT; 206 AA.

AC 08BZE4;
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Truncated MAPK phosphatase 7 homolog.
 GN DEERD213E.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
 RX MEDLINE=22354683; PubMed=1246651;
 RA The FANTOM Consortium.
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60 770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK035652; BAC29138.1; -.
 DR MGD; MG1:1098836; DEERD213E.
 DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.

DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SMO0450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
SQ SEQUENCE 206 AA; 2256 MW; E7FD5C01ADC2F73 CRC64;
Query Match 26.1%; Score 891; DB 11; Length 206;
Best Local Similarity 97.7%; Pred. No. 2.2e-60;
Matches 173; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAHEMIGTQVTERLVALLSGTEKVLIDSRPFVEYNTSHILPAININCSKIMKRRIQQ 60
DB 1 MAHEMIGTQVTERLVALLSGTEKVLIDSRPFVEYNTSHILPAININCSKIMKRRIQQ 60
QY 61 DKVITELIHSANHKVDIDCSQKVVVYDSSQVASSPCFLTVLLGKESFNSVHL 120
DB 61 DKVITELIHSANHKVDIDCSQKVVVYDSSQVASSPCFLTVLLGKESFNSVHL 120
QY 121 LAGGFAERSRCFPGICEKSTLVPTCISQPCLPVANTGPTRIILPNLYLGCORVYLNK 177
DB 121 LAGGFAERSRCFPGICEKSTLVPTCISQPCLPVANTGPTRIILPNLYLGCORVYLNK 177
RESULT 12
Q08ST19 PRELIMINARY; PRT; 657 AA.
ID Q08ST19
AC Q08ST19;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein F08B1.1a.
GN F08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodetidae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Chisoe S.;
RT "The sequence of C. elegans cosmid F08B1.";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR HSSP; Q16828; IMKP.
DR WormPep; F08B1.1a; CE27918.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0017017; F:Map kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:Protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; D:Serine/threonine phosphatase; IEA.
DR InterPro; IPR000343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPc; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SMO0450; DSPc; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.

DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein; Hydrolase.
SQ SEQUENCE 657 AA; 71002 MW; A0D9153DE6326B43 CRC64;
Query Match 17.9%; Score 611; DB 5; Length 657;
Best Local Similarity 29.3%; Pred. No. 4.2e-38;
Matches 201; Conservative 104; Mismatches 198; Indels 184; Gaps 25;
QY 10 IYTERLVALLSGTEKVLIDSRPFVEYNTSHILPAININCSKIMKRRIQQDY----LI 65
DB 9 ISTCGLAALREAPDITLVVDCGFTERYNESHVHSHMNAFSLIRRLFENTLDNCL 68
QY 66 TELIHSAN-----KHKVDIDCSQKVVYDQ-----SSQVDSL-----SDCVLYL 110
DB 69 HQMSSCGSGCTKQDEKLDL-----VLYAEEDKPRGKRIASCNAPESFAKIMRVIRER 122
QY 111 LEKS--FNSHYLLAGGFAERSRCFPGICEKS--TLVPTCISQPC--PVANTGPTRIILP 164
DB 123 LEPTDKFRSYVMLEGGKQPAQYQPOLCESSSEGNTRLPOSLSQPTGD--GITLITP 181
QY 165 NLVYGCORVYLNKELIQONGIGVYLNASVYCPKPDFIPE--SHFLRVVNDSPCEKILPWL 223
DB 182 NIYIGSQIDLDETMDALDISVINIMSWTCPSVCIKEDKIMRIPVNDYQEKLSPTF 241
QY 224 DKSVDFTEKAKASNGCVLVHCLAGISRSATIAIYIMKRMDSLDEAYRFVYKERTISP 283
DB 242 PMAYEFLEKCRRAKCLIHCLAGISRSPTAISYIMRYMKMGSDAYRYVYKERRPSISP 301
QY 284 NFNPLGQLDYER-----KIKNOTGASGPKSKL--HLKRNBP----- 322
DB 302 NFNPMGQLLEYENVLIKDHVLDYNQASRPFRHMDYIGPSDLCPKPKASNCVPPGST 361
QY 323 -----VPAVSEGGQKSETPLSPPCADSATSEAAQRPVHPASVSPVSPSLLEDSP 375
DB 362 HDSSPSPSPSVSGSMASEPETSMAASSSTASA-----PSPMPTSS----- 405
QY 376 LVQALSGHLSDRLSDSNKLRKSFSLD-----KSVTSASMAASLHFFSSSDALEY 429
DB 406 --QOTSSEGTAVANG-----KRMMDLGLPHRPVALGUPSRIGTSV-----AEL 447
QY 430 YKPSITLDGTNKKCPSPVQGLSEQTP---ETSPDKEASIPKKLGTARPSDSQSKRLHS 486
DB 448 PSPSTELSRIS----FNGPAILAPSTPLNFTNFCRNPSTIIV----- 486
QY 487 VRTSSSGTAQRSLSPLRSGSVEDNYHTSFRLGLSTQQHLYKSAGLGLKGMHSDILAP 546
DB 487 --ASSSREVLITLTPPAASSS-----STSS----- 509
QY 547 QTSTPCLTSSWYATSSH-----FYSASAITGSGASYSAGCSQLPTCGDQVYSVRR 600
DB 510 --SEPSFDFSPSSSSSIIVENPPFASSTEVPASSSISTPSGSO----- 553
QY 601 QKPSDRADRSRSMHBSPEKOFKRRS 627
DB 554 STPASASSSAASRCRMKGFKVFSSKKA 580
RESULT 13
Q09GQ52 PRELIMINARY; PRT; 143 AA.
ID Q09GQ52
AC Q09GQ52;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Truncated MAPK phosphatase 7.
GN MKP7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Montpetit A., Boily G., Simmet D.;

RT "A detailed transcriptional map of the chromosome 12p12 tumor suppressor locus."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AY038927; AAK69770.1; -
 DR GO, GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO, GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro, IPR008343; MAPK_phosph.
 DR InterPro, IPR001763; Rhodanese-like.
 DR Pfam, PF00581; Rhodanese.1.
 DR PRINTS, PR01764; MARKPHPTASE.
 DR SMART, SM00450; RHOD.1.
 DR PROSITE, PS50206; RHODANES_3; 1.
 DR SEQUENCE 143 AA; 16077 MW; 5213A213AA7E5974 CRC64;
 SQ
 Query Match 17.8%; Score 610; DB 4; Length 143;
 Best Local Similarity 98.4%; Pred. No. 4.8e-39;
 Matches 144; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MAHEMIGTOIVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKIMKRRLQ 60
 DB 1 MAHEMIGTOIVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKIMKRRLQ 60
 QY 61 DKVLITELIQHSAKHVDIDCSQKVVVYDQSSQDVASLSSDCEFLTVLLGKLEKSFNSVHL 120
 DB 61 DKVLITELIQHSAKHVDIDCSQKVVVYDQSSQDVASLSSDCEFLTVLLGKLEKSFNSVHL 120
 QY 121 LAGGFA 126
 DB 121 LAGADA 126

RESULT 14
 Q8ST18 PRELIMINARY; PRT; 606 AA.
 AC Q8ST18, 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Hypothetical protein F08B1.1b.
 GN F08B1.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Chinese S.;
 RT "The sequence of C. elegans cosmid F08B1.";
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, U31718; AAK68300.1; -
 DR HSSP, Q16828; IMKP.
 DR WormBep, F08B1.1b; CE27919.
 DR GO, GO:0016787; P:hydrolase activity; IEA.
 DR GO, GO:0006138; P:protein tyrosine/serine/threonine phosphata. . .; IEA.
 DR GO, GO:0006701; P:protein amino acid dephosphorylation; IEA.
 DR InterPro, IPR000340; DS_phosphatase.
 DR InterPro, IPR00387; TYR_phosphatase.
 DR Pfam, PF00782; DSpC; 1.

DR SMART, SM00195; DSpC; 1.
 DR PROSITE, PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE, PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE, PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR KW Hypothetical protein; Hydrolase.
 DR SEQUENCE 606 AA; 64962 MW; 5B71795C6BD58058 CRC64;
 SQ
 Query Match 14.6%; Score 497.5; DB 5; Length 606;
 Best Local Similarity 28.9%; Pred. No. 2e-29;
 Matches 160; Conservative 77; Mismatches 159; Indels 157; Gaps 19;
 QY 123 GGFAPSRRCFPGLECEKRS--TLVPTCISQPC--PVANIGPRLIPNLVIGCORVLANKE 178
 DB 86 GGFAPQAOQYPOLCESSEGMTRLPQSLSQPTGD-GITLITPNIYLGSQIDSLDET 144
 QY 179 LIQONGIGVYLAASITCPKPDPIPE-SHFLRPVNDSEFEKILPWLDKSDVDFLEKAKSN 237
 DB 145 MDALDISVIMISMCPKSCVCIKEDKNFMRIPVNDSEYQKLSPPYPMAYEFLEKCRAG 204
 QY 238 GCVLVHCLAGISRSATIAIAYIMKMDSLDAYRPVKEKRPISNPNFLQGLDYEK- 296
 DB 205 KKLCHLCLAGISRSPTLAIYIMRYKMGSDAYRYKERRPSISNPNFQGLLEYENV 264
 QY 297 ----KIKNGTASGPKSKLKL--HLEKNEP-----VPAVSEG 329
 DB 265 LKDHVLDVNGASRRPHNDYGGPSDLCPKVPKASNSCVFPGTHDESSPSSEVSEG 324
 QY 330 GOKSETPLSPPCADSATSEPAQRPVHPASVSVSVQPSLLDEFLVALSGHLASDR 389
 DB 325 SAASEPETSSSAASSSTASA-----PSMSTSE-----CGTSSGTNVANG 366
 QY 390 LEDSNKLKRSFSLDI-----KSVYSASMASLHGFSSEDALEYKSTLLDGTNKC 443
 DB 367 ----KKNMTMDGLPHRPALGLPSRTGTSV-----AELPSTELSRIS---- 407
 QY 444 QSPVQELSEQTP---ETSPDKEASIPKQLQTPAPSDSQSKRLSVRTSGTAQRL 500
 DB 408 -FNGEPAIAPSPILNFTNTPCNSPIIV-----ASSREVIITLP 447
 QY 501 SPLHSGSVEDNYHTSFLGLSTSOQHLTKSAGLGLKWHSDILAPQTSPTSLTSSWYA 560
 DB 448 TPAASSSS-----STS-----SEPFDFSSPFS 470
 QY 561 TESSH-----FYSASAIYGSASAYSQGLPTCGDVSVRRQKSDADSRSH 614
 DB 471 SSSSSTIVENPFPASTEVPAGSSSISTPSGQ-----STPASASSASAKRC 516
 QY 615 EESPEKQPKRRS 627
 DB 517 RMKGFPKVPKSKA 529

RESULT 15
 Q7SZF3 PRELIMINARY; PRT; 367 AA.
 AC Q7SZF3, 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Body;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 21:10:41; Search time 132.037 Seconds
(without alignments)
2794.987 Million cell updates/sec

Title: US-10-029-345A-109
Perfect score: 3418
Sequence: 1 MAHEMGTGTVTERLVALLE.....LKVSGSSFRSSGMEIIEVS 665

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10029345/rnat 21062004 164146 29052/app_query.fasta_1.1294
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPT=0
-LOOPT=0 -INIT=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345 @CNC 1.1 76 @rnat 21062004 164146 29052 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGADEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3406	99.6	1998	4	US-09-816-494-3
2	3406	99.6	3544	4	US-09-816-494-1
3	1326	38.8	2377	4	US-09-920-668-3
4	480	14.0	2109	4	US-09-016-434-1135
5	480	14.0	2109	4	US-09-023-655-946
6	469	13.7	1830	4	US-09-557-921-1
7	453	13.3	1238	4	US-09-702-705-803
8	453	13.3	1238	4	US-09-736-457-803
9	453	13.3	1238	4	US-09-614-124B-803
10	453	13.3	1238	4	US-09-671-325-803
11	453	13.3	1238	4	US-09-589-184-803
12	451.5	13.2	2240	4	US-09-016-434-1100

13	450.5	13.2	2064	4	US-09-702-705-825	Sequence 825, App
14	450.5	13.2	2064	4	US-09-736-457-825	Sequence 825, App
15	450.5	13.2	2064	4	US-09-614-124B-825	Sequence 825, App
16	450.5	13.2	2064	4	US-09-671-325-825	Sequence 825, App
17	450.5	13.2	2064	4	US-09-589-184-825	Sequence 825, App
18	450.5	13.2	2109	4	US-09-702-705-826	Sequence 826, App
19	450.5	13.2	2109	4	US-09-736-457-826	Sequence 826, App
20	450.5	13.2	2109	4	US-09-614-124B-826	Sequence 826, App
21	450.5	13.2	2109	4	US-09-671-325-826	Sequence 826, App
22	450.5	13.2	2109	4	US-09-589-184-826	Sequence 826, App
23	447	13.1	1993	2	US-08-990-379-2	Sequence 2, Appl1
24	445	13.0	1238	2	US-08-530-290-11	Sequence 11, Appl1
25	428.5	12.5	2000	4	US-09-016-434-1291	Sequence 1291, Ap
26	427.5	12.5	2303	4	US-09-922-146-3	Sequence 3, Appl1
27	425	12.4	944	4	US-09-371-671B-10	Sequence 10, Appl
28	405.5	11.9	1987	2	US-08-990-379-1	Sequence 1, Appl1
29	382.5	11.2	1619	4	US-09-702-705-801	Sequence 801, App
30	382.5	11.2	1619	4	US-09-736-457-801	Sequence 801, App
31	382.5	11.2	1619	4	US-09-614-124B-801	Sequence 801, App
32	382.5	11.2	1619	4	US-09-671-325-801	Sequence 801, App
33	382.5	11.2	1619	4	US-09-589-184-801	Sequence 801, App
34	382.5	11.2	4637	4	US-09-702-705-804	Sequence 804, App
35	382.5	11.2	4637	4	US-09-736-457-804	Sequence 804, App
36	382.5	11.2	4637	4	US-09-614-124B-804	Sequence 804, App
37	382.5	11.2	4637	4	US-09-671-325-804	Sequence 804, App
38	382.5	11.2	4637	4	US-09-589-184-804	Sequence 804, App
39	380	11.1	279	4	US-09-016-434-91	Sequence 91, Appl
40	295.5	8.6	539	4	US-09-389-681-311	Sequence 311, App
41	295.5	8.6	539	4	US-09-620-405B-311	Sequence 311, App
42	295.5	8.6	539	4	US-09-339-338-311	Sequence 311, App
43	295.5	8.6	539	4	US-09-433-826B-311	Sequence 311, App
44	295.5	8.6	539	4	US-09-604-287A-311	Sequence 311, App
45	295.5	8.6	539	4	US-09-834-759-311	Sequence 311, App

ALIGNMENTS

RESULT 1
US-09-816-494-3
Sequence 3, Application US/09816494
Patent No. 6664089
GENERAL INFORMATION:
APPLICANT: (Meyers), Rachel A.
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494 - POTSID2
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1998
TYPE: DNA
ORGANISM: Homo sapiens
US-09-816-494-3

POTSID2
NTE SID3

Alignment Scores:			
Pred. No.:	0	Length:	1998
Score:	3406.00	Matches:	663
Percent Similarity:	99.85%	Conservative:	1
Best Local Similarity:	99.70%	Mismatches:	0
Query Match:	99.65%	Indels:	0
DB:	4	Gaps:	0
US-10-029-345A-109 (1-665) x US-09-816-494-3 (1-1998)			
Qy	1	MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu	20
Db	1	ATGGCCCATGATGATGATGGAATTTGTTACTGAGAGTGTGCTGCTGCGGAA	60
Qy	21	SeGClThrcGluYsvaLleuIleapSerArgProPheValGluYrAanThrSer	40

```
Db 61 AGTGAACGAAAAAGCTGCTGTAATTGATACCGGCAATTTGGAAATCAATTCATCC 120
Qy 41 HistileuGlualalAasnllAasnlySerlySleuMetLysArgArgLeuGln 60
Db 121 CACATTTTGAAGCATTAATATCAATGCTTCAAGCTTATGAGCGAAAGTTGCAAG 180
Qy 61 AsplysValleuHleThrGluLeuHleGlnHleSerAlalyHleValAspHleAsp 80
Db 181 GACAAAGTTTAAATACAGAGCTCATCCAGCATTCAGAAACATTAAGGTTGACATTGAT 240
Qy 81 CysSerGlnLysValValValTyzaSpGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 241 TGGAGTCAGAAAGTTGTTAGTTTACATCAAGCTCCAGAGATTTGGCTCTCTCTTCA 300
Qy 101 AspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPheAsnSerValHleLeu 120
Db 301 GACTGTTTTCACAGTACTTCTGGGGTAACTGGAGAAAGACTTCACTCTGTTCACTG 360
Qy 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGlnGlyLysSer 140
Db 361 CTGGAGGTGGGTTTGTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATATCC 420
Qy 141 ThrLeuValProThrCysHleSerGlnProCysLeuProValAlaAsnHleGlyProThr 160
Db 421 ACTGATGCTTACTGCTGATTTCTCAGCTTGTCTTACTGTTGCCAATTTGGGCAACC 480
Qy 161 ArgHleLeuProAsnLeuTyLeuGlnLysGlnArgAspValLeuAsnLysGlnLeuHle 180
Db 481 CGAATTTCTTCCCAATCTTATCTTGGCTGCGCAAGCATGTCTCCACACAGAGCTGATG 540
Qy 181 GlnGlnAsnGlyHleGlyTyTyValLeuAsnAlaSerTyThrCysProLysProAspPhe 200
Db 541 CAGCAGATGGGATTTGTTATGTTTAAATGCCACCAATACCTGTTCCAAAGCTGACTTT 600
Qy 201 HleProGlnSerHleSphLeuArgValProValAsnAspSerPheCysGlnLysHleLeu 220
Db 601 ATCCCCGAGTCTCAATTTCTGCGTGTGCTGAGATGACAGCTTTTGGAAATTTTGG 660
Qy 221 ProThrLeuAspLysSerValAspPheHleGlnLysAlaLysAlaSerAsnGlyCysVal 240
Db 661 CCGGGTTTGAACAAATCGATGATTTCAATTGAAAGCAAAAGCTCCAAATGGATGTGTT 720
Qy 241 LeuValHleCysLeuAlaGlyHleSerArgSerAlaThrHleAlaHleAlaTyHleMet 260
Db 721 CTAATGCACTTTTAACTGAGGATCTCCCGCTCCGCAACATGCTATGCTTACATCATG 780
Qy 261 LysArgMetAspMetSerLeuAspGlnAlaTyArgPheValLysGlnLysArgProThr 280
Db 781 AAGAGATGAGCAATGCTTTTGAATGAACTTACAGATTTTGAAGAAAAAGACCTACT 840
Qy 281 HleSerProAsnPheAsnPheLeuGlnLysLeuAspTyArgLysLysLysHleLysAsn 300
Db 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTCGAGCTATGAGAAACAAATTAAGAAC 900
Qy 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLysLeuHleSleuGlnLysProAsn 320
Db 901 CAGAGTCGAGCATCAGGSCCAAGAGCAAACTCAAGCTGCTGCACTGGAAGAACCAAT 960
Qy 321 GluProValProAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerProPro 340
Db 961 GAACCTGCTGCTGCTCAGAGGTTGAGACAGAAAGACAGAGCCCTCAATCCACCC 1020
Qy 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHleProAlaSerVal 360
Db 1021 TGTGCCCATCTGCTACCTCAGAGCGACAGACAAAGGCCCGGTGCATCCCGCAAGGTG 1080
Qy 361 ProSerValProSerValGlnProSerLeuGlnLysAspSerProLeuValGlnAlaLeu 380
Db 1081 CCCAGCGTCCAGCGTCCAGCGCTGCTGTTAAGAGACAGCCCGCTGCTACAGCGCTC 1140
Qy 381 SerGlyLeuHleSleuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 400
1141 AGTGGCTGACACTGTCCGAGACAGGCTGGAAGACAGCAATTAAGTCAAGCTTCTTC 1200
Qy 401 SerLeuAspHleLysSerValSerTySerAlaSerMetAlaAlaSerLeuHleGlyPhe 420
Db 1201 TCTTGTGATATCAAAATAGTTTCAATATTCAGCCAGCATGAGCATCTTACATGGCTTC 1260
Qy 421 SerSerSerGlnAspAlaLeuGlnTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 440
Db 1261 TCTTCATCAGAGATGCTTTTGAATTAATCAAACTTCCACTACTCTGATGGAGCAAC 1320
Qy 441 LysLeuCysGlnPheSerProValGlnLysLeuSerGlnGlnThrProGlnThrSerPro 460
Db 1321 AAGCTATGACAGTTCTCCCTGTTACAGGACTATCGAGAGACATCCCGAAACCAAGTCT 1380
Qy 461 AsplysGlnGlnAlaSerHleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
Db 1381 GATTAAGAGAGAGCAGCAGATCCCAAGAACTGAGCCGAGCTTTCAGACAGCAG 1440
Qy 481 SerLysArgLeuHleSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db 1441 AGCAGAGATTTGATTTGATTCAGAACCCAGACAGAGTGGACCCGCAAGAGTCCCTTTA 1500
Qy 501 SerProLeuHleArgSerGlySerValGlnAspAsnTyHleThrSerPheLeuPheGly 520
Db 1501 TCTTCATGATCAGATGAGTGGAGCGTGGAGAGCAATTAACACACAGCTTCTTTCGCGC 1560
Qy 521 LeuSerThrSerGlnGlnHleSleuThrLysSerAlaGlyLeuGlnLysGlyTyThrHis 540
Db 1561 CTTTCCACGACGACGACGACCTCACAGATGCTGGCTGGGCTTAAAGGCTGACAC 1620
Qy 541 SerAspHleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyTyPheAla 560
Db 1621 TCGATATCTTGGCCCCCGACGACCTTACCCCTTCCCTGACGAGAGCTGTATTTTGGC 1680
Qy 561 ThrGlnSerSerHleSphTyTySerAlaSerAlaHleTyGlyGlySerAlaSerTySer 580
Db 1681 ACAGAGTCTCAGACATTTCTGCTGCTGCTGACCATCAACGAGCGAGTGCAGTTACTCT 1740
Qy 581 AlaTySerCysSerGlnLeuProThrCysGlyAspGlnAlaTyTySerValArgArgArg 600
Db 1741 GCTTACAGCTGACGACGACGCTCCACTGCGAGAACCAATCTATTTCTGCGCAGGCGG 1800
Qy 601 GlnLysProSerAspArgAlaAspSerArgSerTyHleGlnLysSerProPheGln 620
Db 1801 CAGAGCCAGATGACAGAGCTGACTCGCGCGAGGCTGATGAAGAGACCCCTTTGAA 1860
Qy 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerHleMetSerGlnAsn 640
Db 1861 AAGCAGTTTAAAGCGAGAACCTGCCAAATGGAATTTGAGAGAGCATCATGTCAGAGAAC 1920
Qy 641 ArgSerArgGlnGlnLeuGlnLysValGlySerGlnSerSerPheSerGlySerMetGln 660
Db 1921 AGGTACGGGAAAGAGCTGGGAAAGTGGGAGTCAATGCTTACTTTTGGGACAGATGAA 1980
Qy 661 HleHleGlnValSer 665
Db 1981 ATCATTTAGGCTTCC 1995
RESULT 2
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OR INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OR INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR APPLICATION NUMBER: 2001-03-23
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2583)
US-09-816-494-1

Alignment Scores:
Pred. No.: 0 Length: 3544
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-816-494-1 (1-3544)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB ATGACCCTAAGAGATGATGGAACTCAAAATGTTACTGAGAGTGGCTGCTGCTGGAA 648
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB AGTGGAAAGGAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 708
QY 41 HisIleGluGluAlaIleAsnIleAsnGlySerIleValMetIleValArgLeuGln 60
DB CACATTTTGGAACTTAAATATCACTGCTCCAACTTATGAAAGCGAAGTTGCAACAG 768
QY 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB GACAAAGTGTAAATATACAGAGCTATCAGCAATTCAGCAAACTAAAGTTGACATTGAT 828
QY 81 CysSerGlnIleValValIleValIleValIleValIleValIleValIleValIleVal 100
DB TGCAGTCAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888
QY 101 AspCysPheLeuThrValIleLeuGlyValLeuGlyValLeuGlyValLeuGlyVal 120
DB GACTGTTTCTCACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 948
QY 121 LeuAlaGlyValIleAsnIleAsnGlySerIleValMetIleValArgLeuGln 140
DB CTTCAGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB ACTCTAGTCCCTTACCTGATCTTCTGAGCTTCTGATCTTCTGATCTTCTGATCTTCTG 1068
QY 161 ArgIleLeuProAsnLeuIleGlyValLeuGlyValLeuGlyValLeuGlyValLeu 180
DB CGAATTCCTTCCATCTTATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1128
QY 181 GlnGlnAsnIleIleGlyValLeuAlaAsnIleAsnIleAsnIleAsnIleAsnIle 200
DB 1129 CACACAGATGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1188
QY 201 IleProGlnSerIlePheLeuArgValProValAsnAspSerPheCysGluValIleLeu 220
DB 1189 ATCCCGCAGTCTCATTTCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
QY 221 ProTrpLeuAspIleValIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 240
DB 1249 CCGGTGTGGACAAATCAGTATGATTTTATGAGAAACAAAGCTCCATGATGATGATG 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerLeuAlaIleAlaIleAlaIleAla 260
DB 1309 CTAGTGACAGTGTAGTGGATCTCCGCTCCGCAACATCGCTATCGCTATCATCATG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaIleValArgPheValIleGluValArgProThr 280
DB 261 LysArgMetAspMetSerLeuAspGluAlaIleValArgPheValIleGluValArgProThr 280
```

```
DB 1369 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGTGTAAAGAAAAAGACCTACT 1428
QY 281 IleSerProAsnPheAsnIleGlyIleLeuAspTrpGluIleValIleValIleValIle 300
DB 1429 ATATCTCCAAATCTTAATTTTCTGGGCAATCTCTGACATATGAAAGAAATTAAGAAC 1488
QY 301 GlnThrGlyAlaSerGlyProIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 320
DB 1489 CAGACTGAGACATCAGGCGCAAGAGCAAACTCAACTGCTGCACTGAGAAAGCCAAAT 1548
QY 321 GluProValProAlaValSerGluGlyGlnIleValSerGluThrProLeuSerProPro 340
DB 1549 GAACCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1608
QY 341 CysAlaAspSerLeuThrSerGluAlaIleGlyIleValArgProValHisProAlaSerVal 360
DB 1609 TGTGCGACTGCTCACTCAGAGGAGAGCAAGAGCCGCTGATCCCGCAGCGTG 1668
QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
DB 1669 CCGAGCTGCCAGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1728
QY 381 SerGlyLeuHisIleSerAlaAspArgLeuGluAspSerAlaIleValArgSerPhe 400
DB 1729 AGTGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1788
QY 401 SerLeuAspIleIleValSerValSerIleValSerMetAlaIleSerLeuHisGlyPhe 420
DB 1789 TCTCTGATATCAAAATCAGTTCATATTCAGCAGATGAGCAATCTTACATGCTTC 1848
QY 421 SerSerSerGluAspAlaLeuGluIleValIleValIleValIleValIleValIleVal 440
DB 1849 TCTCTATCAGAGATGCTTTGAAATCTAACAACCTTCACTGCTGATGGAGCAAC 1908
QY 441 LysLeuCysGlnPheSerProValGlnIleValSerGluIleThrProGluThrSerPro 460
DB 1909 AAGCTATGCAAGTTCCTCCCTGTTAGAACTATCGAAGAGACTCCCGAAACAGTCT 1968
QY 461 AspIleGluGluAlaSerIleProIleValIleValIleValIleValIleValIleVal 480
DB 1969 GATTAAGAGAGAACCCAGATCCCAAGAGCTGCAACCGCGCTTCAGACAGCAG 2028
QY 481 SerIleArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
DB 2029 AGCAGAGATGATTCGATCGATCAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2088
QY 501 SerProLeuHisArgSerGlySerValGluAspAsnIleThrSerPheLeuPheGly 520
DB 2089 TCTCCACTGATGAGATGGAGAGCGTGAAGCAATTAACAACAGCTTCCTTTCCGCG 2148
QY 521 LeuSerThrSerGlnGlnIleValThrIleValSerAlaGlyIleValGlyIleValIle 540
DB 2149 CTTTCCACAGCAGCAGACAGACCTCAGAAATGCTGCGCTTGAAGGCTGGCAC 2208
QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerIleThrSerSerTrpIlePheAla 560
DB 2209 TCGGATATCTTGGCCCCCAGACCTTACCTCTCCCTGACACAGAGCTGATTTTGGC 2268
QY 561 ThrGlnSerSerIlePheTrpSerAlaSerAlaIleTrpGlyIleSerAlaSerTrpSer 580
DB 2269 ACAGAGTCCCTCACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2328
QY 581 AlaTrpSerCysSerGlnIleProThrCysGlyAspGlnValIleValIleValIleVal 600
DB 2329 GCTTACAGCTGACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2388
QY 601 GlnIleProSerAspArgAlaAspSerArgSerTrpHisGluGluIleValSerProPheGlu 620
DB 2389 CAGAACCCAGATGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2448
QY 621 LysGlnPheValArgArgSerCysGlnMetGluPheGlyIleValSerIleMetSerGluAsn 640
DB 2449 AAGCAGTTTAAAGCAGAGAGCTCCAAATGAAATTTGAGAGAGCATGTCATGACAGAAC 2508
```

```
QY 641 ArgSerArgGluGluLeuGluValGlySerGlnSerSerPheSerGlySerMetGlu 660
DB 2509 AGGTACGGGAAAGAGCTGGGAAAGTGGGAGTCACTGCTTTTGGGGCAGCATGGAA 2568
QY 661 IleIleGluValSer 665
DB 2569 ATCATGAGGTCTCC 2583

RESULT 3
US-09-920-668-3
; Sequence 3, Application US/0920668
; Patent No. 6482644
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowsett
; APPLICANT: Brett P. Monia
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 8 EXPRESSION
; FILE REFERENCE: RTS-0246
; CURRENT APPLICATION NUMBER: US/09/920,668
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)...(2012)
US-09-920-668-3

Alignment Scores:
Pred. No.: 1,34e-128 Length: 2377
Score: 1326.00 Matches: 314
Percent Similarity: 58.54% Conservative: 94
Best Local Similarity: 45.05% Mismatches: 185
Query Match: 38.79% Indels: 104
Gaps: 21

US-10-029-345A-109 (1-665) x US-09-920-668-3 (1-2377)
QY 1 MetaIaHISgluMetIleGlyThrGlnIleVal---ThrGluArgLeuValAlaLeuLeu 19
DB 135 ATGGCTGGAGACGGGCTCCGAGGAAGGTGATGTCAGCAAGCTGGCCAGCTGTGTG 194
QY 20 GluSerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 39
DB 195 CGGGCGGGCGCTGGGGGGCGCTGTGATTCAGACAGCCCTCTCTGTTGAGATACAGAC 254
QY 40 SerHisIleLeuGluAlaIleAsnIleAsnGlySerIleLeuMetIleValArgArgLeuGln 59
DB 255 TGGCATGTGCTCAGCTCCGTCAACATCTGCTGCTCCAGCTGTGTAAGCGGGCTGAG 314
QY 60 GlnAspIleValLeuIleThrGluLeuIleGlnHisSerAlaValSHsIleValAspIle 79
DB 315 CAGGGCAAGGTGACCATTCGAGCTCATTCAGCGGGCTGACACGAGCAAGCTGAGAGCT 374
QY 80 AspCysSerGlnIleValIleValIleValIleValIleValIleValIleValIleVal 99
DB 375 ACGGAGCCACAGAGAGTGTGTCTATACACAGACAGCGGGAGCGCCAGGTCTGAGCC 434
QY 100 SerAspCysPheLeuThrValLeuLeuGluIleValLeuGluIleValSerPheAsnSerValHis 119
DB 435 GCAAGCAGAGCTTCCTCATCTGCTGCTGAGCAAGCTGAGCGGTCTTCAACAGCGTGGCC 494
QY 120 LeuLeuAlaGlyValPheAlaGluPheSerArgCysPheProGlyLeuGluGlyLys 139
DB 495 ATCTCTACTGGGGGCTTGCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 554
QY 140 ---SerThrLeuValProThrCysAlleSerGlnProCysLeuProValAlaAsnIleGly 158
DB 555 CCGTCTGCGCTGACCATGAGCTCTCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 614
QY 159 ProThrArgIleLeuProAsnLeuThrLeuGluCysGlnArgPheValLeuAsnLysGlu 178
```

```
DB 615 CTGACCCGATCTCTGCTCATCTTACCTGGCTCGGAGGAAGAGCTCTTAACAAGGAT 674
QY 179 LeuIleGlnIleAsnGlyIleGlyThrValLeuAsnAlaSerThrCysProLysPro 198
DB 675 CTGATGACGCAAAATGGAATAGCTACGTCCTCAAGCCAGCAACCTCTGCCCAAGCCT 734
QY 199 AspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLys 218
DB 735 GACTTCATCTGCGAGACCCGCTTCATGGGATCGGATCAACAGCAACATCTGTGAAAAA 794
QY 219 IleLeuProThrLeuAspLysSerValAspPheIleGluIleValAlaValSerAspGly 238
DB 795 CTGCTGCCCTCGCTGGAGCAAGTCAATGATTCATGATTAAGCCAGCTCTCAGCTGC 854
QY 239 CysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyr 258
DB 855 CAAGTCACTGTCACCTGTCTGTGGCTGATCTCCGGCTCTGCCAATCGCCATCGCTTAC 914
QY 259 IleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGlyLysArg 278
DB 915 ATCATGAAGACCAATGGGCATGTCTCCGACAGCGCTTCAAGGTTCTGGAAGACAGCGC 974
QY 279 ProThrIleSerProAsnPheAsnPheLeuGlyIleLeuLeuAspIleValLysIle 298
DB 975 CCGTCCATCTCGCCCACTTCAACTTCTGTGGCCAGCTGTGAGATCAAGCCAGCGCTG 1034
QY 299 LysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLeuHisLeuGluLys 318
DB 1035 AAGCTGCTGGCCGCTCGACAGCGCCGAGC-----ACC 1070
QY 319 ProAsn-----GluProValProAlaValSerGluGlyGlnLysSerGluThr 335
DB 1071 CCTTCAGGAGCGCGGAGCTCCGCGAGTCTGCGCGGGGCCCGCTGCCAGCGCTG 1130
QY 336 ProLeuSerProProCysAlaAspSerAlaThr-----SerGluAlaIleGlyArg 353
DB 1131 CCA-----CCACTTACTGAGAGAGCGCTGCCACAGGAAATCGGCTGCGAGGAGGCG 1184
QY 354 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuLeuGluAsp 373
DB 1185 GGCTGTGCGCGCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1235
QY 374 SerProValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspProLeuGluAspSer 393
DB 1236 AGCGACCTGAGAGAGGCTGCGCGCTGCACTTCTCTGAGCGCGCTGAGAGGAGCT 1295
QY 394 AsnLysLeuLysArgSerPheSerLeuAspIleLysSerValSerThrSerAlaSerMet 413
DB 1296 AACCGCTCAAGGCTCTCTTCTCTGAGCATCAAGTCTGCC----- 1337
QY 414 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGluIleTyrThrLysProSer 433
DB 1338 -----TACCCTCTTACC 1349
QY 434 ThrThrLeuAspGlyThrAsn-----LysLeuCysGlnPhe 445
DB 1350 AGCGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1409
QY 446 ---SerPro-----ValGlnLeuSerGluGlnThrProGluThrSerProAsp 461
DB 1410 GACAGCCCGGTGGGGGCGCGCTGGGCTGTCTGTCGCGCCAGCGCGAGC---AGCCCGAGC 1466
QY 462 LysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGlnSer 481
DB 1467 GCGCGCGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1511
QY 482 LysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuSer 501
DB 1512 -----CCGCGCGCTCTCCCGCGCGC-----TTC 1535
QY 502 ProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly--- 520
```



```
DB 1536 CCGCGCAGACCTCGGC-----CTGACCTCGGCAT 1568
QY 521 -----LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuGly 538
DB 1569 GCGCGCCGCGACCTCGCGGACAGCGGCTCTCGGCTCGGCGCGCGCGCGCGCGCG 1628
QY 539 -----TrpHisSerAspIleLeuAlaProInThrSer 549
DB 1629 CCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1682
QY 550 ThrProSerLeuThrSerSerTrpTyPheAlaThrGluSerSerHisPheTySerAla 569
DB 1683 AGCCGCTCGCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1721
QY 570 SerAlaIleTyGlySerAlaSerTySerAlaTySerTySerGlnLeuProThr 589
DB 1722 GCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1781
QY 590 CysGlyAspGlnValTySerValArgArgGlnLysProSerAspArgAlaAspSer 609
DB 1782 CCAAGCGCGCGCGACGACCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1841
QY 610 ArgArgSerTrpHisGlnGluSerProPheGlnLysGlnPheLysArgArgSerCysGln 629
DB 1842 CGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1901
QY 630 MetGluPheGlyGluSerIleMetSerGluAsnArgSerArg---GluGluLeuGlyLys 648
DB 1902 ATGAGAGTTCAAGAGAGGCG---ATGATGAGAGGCGCGCGCGCGCGCGCGCGCG 1958
QY 649 ValGlySerGlnSerPheSerGlySerMetGluIleIleGluValSer 665
DB 1959 CTGGGCAAGACGAGGAGCTTCTCGGCGAGCGCTGAGGTCAATCAAGAGTGTCC 2009
```

RESULT 4

```
US-09-016-434-1135
; Sequence 1135 Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Sellhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.434
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1135:
; SEQUENCE CHARACTERISTICS:
```

```

; LENGTH: 2109 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 91418933
; US-09-016-434-1135

Alignment Scores:
Pred. No.: 8,466-40 Length: 2109
Score: 480.00 Matches: 121
Percent Similarity: 51.62% Conservative: 70
Best Local Similarity: 32.70% Mismatches: 126
Query Match: 14.04% Indels: 53
DB: 4 Gaps: 9

US-10-029-345A-109 (1-665) x US-09-016-434-1135 (1-2109)
QY 19 LeuGluSerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyPhe 38
DB 430 CTGGAGCTGGCGACACGAGCGGCTGCTGATGAGACTGCGCGCGACGAGACTATACGAG 489
QY 39 ThrSerHisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeu 58
DB 490 TCCTGCACATCGAGTCGCGGCATCACTGCGCATCCCGGCATATCTCGCGCGCGCTG 549
QY 59 GlnGlnAspLysValLeuIleThrGluLeuIleGlnHisSerAla---LysHisLysVal 77
DB 550 CAGAAAGGTAACTCGCGCGCGCGCGCTTTCACCGCGCGCGAGGACCGCGACCGCTTC 609
QY 78 AspIleAspCys---SerGlnLysValValValTyArgGlnSerSerGlnAsp---Val 95
DB 610 ACCCGCGCTGTGGACCGACACAGAGTGCTCTTACGACGAGACGACGACGAGTGAAC 669
QY 96 AlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPhe 115
DB 670 GAGAAATACGCGCGCGCGAGCTGCTGCTGCGGCTGCTCTCAAGAGCTCAAGACGAGGCG 729
QY 116 AsnSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeu 135
DB 730 TGCGCGCGCTTCACTGAGAGGTGCTTCAGAAATTCACAGCCGAGTTCTCCCTGCAT 789
QY 136 CysGluGlyLysSerThrLeuValProThrCys---IleSerGlnProCysLeuProVal 154
DB 790 TGGAG---ACCAATTAAGCGGCTGTGAGCAGACGCTCGCGCGCTTGCAAGTG 843
QY 155 AlaAsnIleGly----- 158
DB 844 CTGGGCTCGGGGCGCTCGGATCACTGACTCTTCTCGGACATCGAGTGTGACTT 903
QY 158 ----- 158
DB 904 GACCGAGACCCCAATATGCAACAGACGCTGATGATCCGCTGCCAAGCAGCGCT 963
QY 159 -----ProThrArgIleLeuProAsnLeuTyTrpLeuGlyCysGlnArgAspValLeuAsn 176
DB 964 TCCTTCCAGTGGAGATCTTCCCTTCTTACCTTGGGCTGTGCCAAGACCTCCACCAAC 1023
QY 177 LysGluLeuIleGlnAsnGlyIleGlyTyValLeuAsnAlaSerTyThrCysPro 196
DB 1024 TTGAGACGTGGAGAAATTCGCGATCAAGTACATCTTAAGCTCACCCCAATTTGCC 1083
QY 197 LysPro---AspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPhe 215
DB 1084 AATCTCTTGAGAACGACGAGAGGTTAATTAATCAACAATCCCATCTCGATCACTGG 1143
QY 216 CysGluLysIleLeuProThrLeuAspLysSerValAspPheIleGluLysAlaLysAla 235
DB 1144 AGCAAAACCTGCTCCAGTTTCTTCCGTGAGGCCATTCTTCAATATGAAGCCCGGCG 1203
QY 236 SerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAla 255
```

```

Db      1204  AAGAACTGTGTGTCTTGTAACATTCGTTGGCTGGCATTAAGCCGCTCACTGCTACT 1265
Qy      256  ILEAIAIYRIIWEETLYSAAGMETASPMETSERLEUASPOLUALATYARPhVEAllys 275
Db      1264  GTGGCTTACCTTATGACGAAGCTCAATCTGTGCATGAAACGATGCCATATACATTGTCAA 1323
Qy      276  GILUYSARGPProthrlIeSerProaAnPhaLeuGlyGluLeuLeuaspYrGlu 295
Db      1324  ATGAAAAAATCCAAACATATCCCTTAACCTTACATTATGAGTACAGTGTGGACATTGAG 1383
Qy      296  LyeLysIleLysAsnGlnThrGlyAlaSerGlyPro-----LysSerLysLeuLys 312
Db      1384  AGAGACGCTG-----GGACTCAGACAGCCATGTCAGACAGAGG99TTCAGACAG 1431
Qy      313  LeuLeuHISreuGluLysProAsnGluProValProAlaValSerGluGlyGluLys 332
Db      1432  CAGCTGATTTTACCAACCTTCCTTCAACCAAGAAATGATACAGAGTGGACTTCTGCATCT 1491
Qy      333  Ser-GluThrProLeuSerProProCys 341
Db      1492  ACGTGAAGACCCCAACCCCTCTCTTGC 1519

RESULT 5
US-09-023-655-946
Sequence 946, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Selhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023.655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 946:
SEQUENCE CHARACTERISTICS:
LENGTH: 2109 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1418933
US-09-023-655-946

Alignment Scores:      8.46e-40      Length:      2109
Pred. No.:

```

Score:	480.00	Matches:	121
Percent Similarity:	51.62%	Conservative:	70
Best Local Similarity:	32.70%	Mismatches:	126
Query Match:	14.04%	Indels:	53
DB:	4	Gaps:	9
US-10-029-345A-109 (1-665) x US-09-023-655-946 (1-2109)			
QY	19 LeuGIuSerGIyThrGIuIysValLeuLeuIIAspSerArgProPheValGIuTYrAsn	38	
Db	430 CTGAGCTGGGCAACAGCGCGCTGCTGATGATGATCGCGCGGAGAGTATACGAG	489	
QY	39 ThrSerHisIIeLeuGIuIIaIIeAsnIIeAsnCysSerIysLeuMetIysArgValLeu	58	
Db	490 TCGTCCACATGAGTGGCGGCATCAACGTGGCCATCCGGGCAATCATGCTGGCGCGCTG	549	
QY	59 GIuGlnAspIysValLeuIIeThrGIuLeuIIeGlnHisSerAla---LysHisIysVal	77	
Db	550 CAAAGAGGTAACTCCGCGCGGCGCGCTTCACGCGCGGCGAGGACGGAGCGGCGCTTC	609	
QY	78 AspIIeAspCys---SerGlnIysValIValTYrAspGlnSerSerGlnPro---Val	95	
Db	610 ACCCGCGCTGGTGGCACCCACACAGTGGTGGCTTACAGAGAGGACGACGACCTGGAAC	669	
QY	96 AlaSerLeuSerSerArgCysPheLeuThrValLeuLeuGIyLysLeuGIyLysSerPhe	115	
Db	670 GAGAAATACGGCGCGGCGAGTCTGTGGCTCGGCGCTGCTCAAGAGCTCAAGAGCGAGGCG	729	
QY	116 AasnSerValHisLeuLeuIIaGIyGIyPheIIaGIuPheSerArgCysPheProGIyLeu	135	
Db	730 TCCCGGCGCTTACTCGGAGAGGTGGCTCAATGATTCGAACCGGAGTTCTCCCTGCAT	789	
QY	136 CysGIuGIyLysSerThrLeuValProThrCys---IleSerGlnProCysLeuProVal	154	
Db	790 TCGGAG-----ACCAATCTAGACGCGCTGGTACAGAGAGTCCGCCCTTGGCCAGTG	843	
QY	155 AlaAsnIIeGIy-----	158	
Db	844 CTGGGCGTCCGGGCGGCTCGGATCAGCTGTGACTTCTCCGACATCGAGCTGACCTT	903	
QY	158 -----	158	
Db	904 GACGAGACCCCAATAGTGCACAAAGACCTCGGATGGTACTCCGCTCTCAACAGCCAGCTT	963	
QY	159 -----ProThrArgIIeLeuProAsnLeuTYrLeuGIyCysGlnArgAspValLeuAsn	176	
Db	964 TCTCTTCCAGTGGAGATCTTCCCTCTCTACTTGGGCTGGCCAAAGATCCACCCAAC	1023	
QY	177 LysGIuLeuIIeGlnGlnAsnGIyIIeGIyTYrValLeuAsnAlaSerTYrThCysPro	196	
Db	1024 TTGAGACGTGTGGAGAAATTGGGATCTCAAGTACATCTTGAACTGCACCCCAATTGGCG	1083	
QY	197 LysPro---AspPheIIeProGlnSerHisIAspLeuArgValIProValAsnAspSerPhe	215	
Db	1084 AATCTCTTTGAGACGACGAGAGATTAAATACAAAGAAATCCCACTCTGGATCACTGG	1143	
QY	216 CysGIuIysIIeLeuProTYrPheAspIysSerValAspPheIIeGIyLysAlaIysAla	235	
Db	1144 AGCCAAACCTGTCCAGTTTTTCTCGAGGCCATTTCTTTCATAGATGAAAGCCGGGCGC	1203	
QY	236 SerAsnGIyCysValLeuValHisCysLeuAlaGIyIleSerArgSerAlaThIIeAla	255	
Db	1204 AAGAACTGGTGGTCTTGGTACATGCTTGGCTGGCATTAAGCCGCTAGTCACTGTACT	1263	
QY	256 IIeAlaIyTIIeMetIysArgMetAspMetSerIleuAspGlnAlaTYrArgPheValIys	275	
Db	1264 GTGGCTTAATCTTATGCAAGAGCTCAACTGTGCATGAAACGATGCTATGACATGTCAAA	1323	
QY	276 GIuIysArgProThrIleSerProAsnPheAsnPheLeuGIyGlnLeuLeuAspTYrGIu	295	
Db	1324 ATGAAAAAATCCAAATATCCCTTAACCTTAACCTTAATGAGGTGAGCTGGACCTTGAG	1383	
QY	296 LysIysIIeLysAsnGlnThrGIyAlaSerGIyPro-----LysSerIleLys	312	

```
Db 1384 AGAGCGGTG-----GAGCTGAGCAGCCCATGTGACAAACAGGGTTCAGCAG 1431
Qy 313 LeuLeuHisLeuGluYsProAngLupProValProAlaValSerGluGlyGlnLys 332
Db 1432 CAGCTGATTTTACACCCCTTCCACCAAGATGTATACAGGTGAGCTCTGTGCATCT 1491
Qy 333 Ser-GluThrProLeuSerProProCys 341
Db 1492 ACGTAAAGACCCACACCCCTCTGTC 1519

RESULT 6
US-09-557-921-1
; Sequence 1, Application US/09557921
; Patent No. 6531810
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125 416
; CURRENT FILING DATE: US/09/557, 921
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-557-921-1

Alignment Scores:
Pred. No.: 9,28e-39 Length: 1830
Score: 469.00 Matches: 107
Percent Similarity: 57.72% Conservative: 65
Best Local Similarity: 35.91% Mismatches: 96
Query Match: 13.72% Indels: 30
Gaps: 7

US-10-029-345A-109 (1-665) x US-09-557-921-1 (1-1830)
Qy 27 LeuLeuHisLeuSerArgProPheValGluThrAnthrSerHisLeuGluAlaIle 46
Db 603 GTATCATTTGACGAGCGCCCTTATGAGTAAACAAGATCATTCAGAGAGCTGTC 662
Qy 47 AsnIleAnCySer---LysLeuMetLysArgArgLeuGlnGlnAspLysValLeuIle 65
Db 663 CACATTAACTGCTGCGATGATCAGCGCGGAGACTGCGAGCAGCGCAAGATCATCTGC 722
Qy 66 ThreGluLeuLeuGln---HisSerAlaLysHisLysValAspLysSerGlnLys 84
Db 723 CTAGACTTGATTTCTCTGTAAGGAAGCAAGACTCTTCAAGAGATCTTTCCAAAGAA 782
Qy 85 ValValValLysArgGlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeu 104
Db 783 ATTATATTTTATGATGAGATACCAATGACCAAGCCGAGTGAATGCCCTCCAGCCACTT 842
Qy 105 ThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeuLeuAlaGly 124
Db 843 CACATAGCTCTCGAGCTCCGTGAAGAGAAAGCAAAAGAACCTCTGCTTGAAGGTGGA 902
Qy 125 PheAlaGluPheSerArgCysPheProGlyLeuGluGluGlyLys----- 139
Db 903 CTTAGTGTGTTTAAACAGAAACATGAAACCTCTGTACCACTCCCTCCAGCTCCAAAG 962
Qy 140 -----SerThrLeuValProThrCysIleSer 148
Db 963 TGCCGGAGGTGGGGGGGGCGCATCGCGGCTCGAGCTTGCACTTCAGCCCATC--- 1019
Qy 149 GlnProCysLeuPro---ValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyr 167
Db 1020 ---CCCAACCAACCTGATGAGACGCTGAGCTCAACCCCATCTTGCCCTCTGTTTC 1076
Qy 168 LeuGlyCysGlnArgAspValLeuAsnLysGluLeuLeuGlnGlnAsnGlyIleGlyTyr 187
```

```
Db 1077 CTTCGCAATGAGCAGGATGCTCAGACCTGAGACACCATGACGCGGTGAACATCGGCTAC 1136
Qy 188 ValLeuAsnAlaSer-----TyrThrCysProLysProAspPheIlePro 202
Db 1137 GTCATACAGTCAACCACTCATCTTCCCTTACCACTATGAGAAAGCCCTGTC----- 1190
Qy 203 GluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeuProTrp 222
Db 1191 -----AACTACAGAGCGGCTGCGCAGCACTGACAGCAACAGCAACCTGCGCACTAC 1244
Qy 223 LeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAnGlyCysValLeuVal 242
Db 1245 TTGTGAAGAGCTTTGATGTTTATGAGAGAGCTCACAGATGTGGAGAGGCGCTTCATC 1304
Qy 243 HisCysLeuAlaGlyLysSerArgSerAlaThrIleAlaIleAlaTyrIleMetLysArg 262
Db 1305 CACTGCAGGCTGGGGGTGTCCTCCGTCGACCATGCTATCCCTTAATGATGAAGCAC 1364
Qy 263 MetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThrIleSer 282
Db 1365 ACTCGATGACCACTGACTGATGCTTAATTAATTGTCAAGCAACCAACCAATTATCTCC 1424
Qy 283 ProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
Db 1425 CCAACTTAATCTTCAATGGGCGAGTGTGATGAGTTGAGAGAACCTTAACAC 1478

RESULT 7
US-09-702-705-803
; Sequence 803, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongrong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Derrick
; APPLICANT: Reiter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121 478C14
; CURRENT FILING DATE: US/09/702, 705
; NUMBER OF SEQ ID NOS: 2000-10-30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 803
; LENGTH: 1238
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-803

Alignment Scores:
Pred. No.: 2.14e-37 Length: 1228
Score: 453.00 Matches: 120
Percent Similarity: 50.52% Conservative: 73
Best Local Similarity: 31.41% Mismatches: 143
Query Match: 13.25% Indels: 46
Gaps: 13

US-10-029-345A-109 (1-665) x US-09-702-705-803 (1-1238)
Qy 19 LeuGluSerGlyThrGluLysValLeuLeuLeuAspSerArgProPheValGluTyrAsn 38
Db 168 CTCGCGAGGCGC---GCAAGATGCTGCTGCTGCTGCACTGCAAGACCTTCTCGGCGCAGC 224
Qy 39 ThrSerHisIleLeuGluAlaIleAsnIleAnCySerLysLeuMetLysArgAlaGlu 58
Db 225 GCGGCTACACTCTTAGTGTGCAACGTGCGCTGTAAACCACTTCGCGCGCGCG--- 281
Qy 59 GlnGluAspLysValLeuIleThrGluLeuIle-----GlnHisSerAlaLysHisLys 76
```

```

Db      282 GCTAAGGCTCGGTGAGCGAGCATCTCTGCCCCCGAGAGAGATACGGGCCG 341
Qy      77 ValAspIleAspCysSerGlnIlyValValIlyAspGlnSerSerGlnAspValAla 96
Db      342 TTGGCTCCGGCCCTTACTGCGCGGTGCTGCTACGACGACGCGCGCGCGCGAG 401
Qy      97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyIlySleuGlnIlySerPhe 115
Db      402 AGCTTCGCGAGACGACGCGTGCCTGGTGGTGGCGCGCGCGCGCGCGAG 461
Qy      116 ---AsnSerValHisLeuLeuAlaGlyIlyPheAlaGlnPheSerArgCysPheProGly 134
Db      462 CGCACCAGCATCTGCTCTCAAAAGCGGCTATAGAGGTTTCTCTCGAGTACCCGAGA 521
Qy      135 LeuCysGlnGlyIlySerThrIleu-----ValProThrCysIleSer 148
Db      522 TTCTGTTCTAAACCAAGGCCCTGGACACCATCCACCCCGGTTCCCGCAGCGCCACA 581
Qy      149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db      582 GAGCCCTTGAACCTGAGACTGAGCTCTGTGGAGACCCACATACACAGACGAGGGTCT 641
Qy      160 ThrArgIleLeuProAsnLeuIlyLeuGlyCysGlnArgAspValLeuAsnIlyLeu 179
Db      642 GTGAGAGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 701
Qy      180 IleGlnIleAsnGlyIlyIleGlyIlyValLeuAsnAlaSerIlyThrCysProIlyProAsp 199
Db      702 CTGAGACCCCTGGGCGATCACGGCTGTGTGTAATGTCTCTCGGACGCGCCAAAC---CAC 758
Qy      200 PheIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIlySle 219
Db      759 TTTCAGAGACACATCATGTAACAGTCCGAGTGAAGAAATAACACAGCGCCGACATC 818
Qy      220 LeuProIlyLeuAspIlySerValAspPheIleGlyIlyAlaIlyAlaSerAsnIlyCys 239
Db      819 AGCTCTCGGTTCATGAGACCATATAGATCATGATCCGTGAAGACCTCCGTGGGCGC 878
Qy      240 ValLeuValHisCysLeuAlaGlyIlyIleSerArgSerAlaThrIleAlaIleAlaIlyIle 259
Db      879 GTCTGTGTCACATCGACGCGGCGATCTGCGGTGCGCCACCATCTGCGCTGACCTG 938
Qy      260 MetIlyArgMetAspMetSerLeuAspGlnAlaIlyArgPheValIlyGlnIlyLeuArgPro 279
Db      939 ATGATGAAGAAACGGGTGAGGCTGAGAGAGGCTTCGAGTTCTGTTAACGACGCGCGACG 998
Qy      280 ThrIleSerProAsnPheAsnPheLeuGlyIlyLeuLeuAspIlyArgIlyIlySleIlys 299
Db      999 ATCATCTCGCCCAACTTCAGCTTCATGGGCGAGCTGCGAGTTCCAGTCCAGGTCTG 1058
Qy      300 AsnGlnThrGlyAlaSerGlyIlyProIlySerIlySleuIlySleuGlnIlyPro 319
Db      1059 GCCACGCTCTGTGCGGAGGCTGCTAGC-----CCC 1091
Qy      320 AsnGlnProValProAlaValSerGlnIlyGlnIlySerGlnIlyProIlyProLeuSerPro 339
Db      1092 TCGGAGCCCTG-----GGGAGCGCGGCAAGACCCCGCCACCCCGC 1133
Qy      340 ProCysAlaAspSerAlaThrSerGlnAlaAlaGlyIlyGlnArgProValHisProAlaSer 359
Db      1134 -----ACCTTCGAGTCTGCTTTCAGCTTTCGCGGCT---TCCGTGGCGC 1172
Qy      360 ValProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAla 379
Db      1173 GTGAGACTCGGCGCCCGACGAGCTGCTTACCTG-----CACAGCCCATACACCTCT 1226
Qy      380 LeuSer 381
Db      1227 CCCAGC 1232

```

RESULT 8
US-09-736-457-803

```

/ Sequence 803, Application US/09736457
/ Patent No. 6509448
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedrick, Tom
/ APPLICANT: Carter, Patrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Lijun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C15
/ CURRENT APPLICATION NUMBER: US/09/736,457
/ NUMBER OF SEQ ID NOS: 1864
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 803
/ LENGTH: 1238
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ US-09-736-457-803

```

```

Alignment Scores:
Pred. No.: 2,14e-37 Length: 1238
Score: 453.00 Matches: 120
Percent Similarity: 50.52% Conservative: 73
Best Local Similarity: 31.41% Mismatches: 143
Query Match: 13.25% Indels: 46
Gaps: 13

```

US-10-029-345a-109 (1-665) x US-09-736-457-803 (1-1238)

```

Qy      19 LeuGlnSerGlyIlyThrGlnIlyValLeuLeuIleAspSerArgProPheValGlnIlyArg 38
Db      168 CTGCGCGAGCGC---GGCAAGTCTGCTGCTGAGACTGACACACGTTCTCGGCGACAGC 224
Qy      39 ThrSerHisIleLeuGlnAlaIleAsnIleAsnCysSerIlySleuMetIlyArgArgLeu 58
Db      225 GCGGCTCATCTCTGATGTTGCTGTCGATCAAGTGGCTGTAAACCATCTGCGCGCGG--- 281
Qy      59 GlnGlnAspIlyValLeuIleThrGlnIle-----GlnHisSerAlaIlyHisIlys 76
Db      282 GCTAAGGCTCGGTGAGGCTGAGACGATCTGCGCCGAGGAGAGAGTACCGCGCGC 341
Qy      77 ValAspIleAspCysSerGlnIlyValValIlyArgPheGlnSerSerGlnAspValAla 96
Db      342 TTGGCTCCGGCCCTTACTGCGCGGTGATGCTATACGACGAGCGCGCGCGCGCGAG 401
Qy      97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyIlySleuGlnIlySerPhe--- 115
Db      402 AGCTTCGCGAGACGACGCGTGCCTGGTGGTGGCGCGCGCGCGCGCGAG 461
Qy      116 ---AsnSerValHisLeuLeuAlaGlyIlyPheAlaGlnPheSerArgCysPheProGly 134
Db      462 CGCACCAGCATCTGCTCTCAAAAGCGGCTATAGAGGTTTCTCTCGAGTACCCGAGA 521
Qy      135 LeuCysGlnGlyIlySerThrIleu-----ValProThrCysIleSer 148
Db      522 TTCTGTTCTAAACCAAGGCCCTGGACACCATCCACCCCGGTTCCCGCAGCGCCACA 581
Qy      149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db      582 GAGCCCTTGAACCTGAGACTGAGCTCTGTGGAGACCCACATACAGACGAGGAGTCTCT 641
Qy      160 ThrArgIleLeuProAsnLeuIlyLeuGlyCysGlnArgAspValLeuAsnIlyLeu 179
Db      642 GTGAGAGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 701
Qy      180 IleGlnIleAsnGlyIlyIleGlyIlyValLeuAsnAlaSerIlyThrCysProIlyProAsp 199

```

```

Db      702 CTGAGAGCCCGGACATCAAGGCTCTGTGATGTCCTCGAGACTGCCAAAC---CAC 758
Qy      200 PheIIProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIle 219
Db      759 TTTGAAGACACTATCACTGACAGTGCATCCAGTGAAGATTAACCAAGCCGACATC 818
Qy      220 LeuProTrieLeuAspLysSerValAspPheIIleGluValAlaValAsnAngIyCys 239
Db      819 AGCTCTGTTGATGAGAGCCATAGAGTACATCATCCCTGAAGAGACTCCGCGGCGC 878
Qy      240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIle 259
Db      879 GTGCTGTGCACTGCCAGGCGGCGCATCTCGCGGTCCGACCAACATCTCCCTGACTACTG 938
Qy      260 MetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIlyGluValArgPro 279
Db      939 ATGATGAAGAAACCGGAGGCTGAGGCTGAGAGAGGCTTCGAGTTCGTTAAGCAGCGCCGAGC 998
Qy      280 ThrIleSerProAsnPheAsnPheLeuGlyGluLeuAspTyrGluValIleLys 299
Db      999 ATCATCTCGCCCAACTTCAGCTTCATGAGGAGCTGCTGAGATTCGAGTCCCAAGTCTG 1058
Qy      300 AsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysIleLeuGluLysPro 319
Db      1059 GCCACGCTCTGTGCTGCGGAGGCTGCTAGC-----CCC 1091
Qy      320 AsnGluProValProAlaValSerGluGlyGluLysSerGluThrProLeuSerPro 339
Db      1092 TCGGAGACCCCTG-----GGGAGCGGGGCAAGACCCCGCCAGCCACC 1133
Qy      340 ProCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSer 359
Db      1134 -----ACCTCGCAAGTCTGCTTTCAGCTTTCGCGTC---TCGTTGGGC 1172
Qy      360 ValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuAlaGlnAla 379
Db      1173 GTGACTTCGGCCCCCAGCAGCTGCTTACTG-----CACAGCCCATATCACCACTCT 1226
Qy      380 LeuSer 381
Db      1227 CCCGAGC 1232

RESULT 9
US-09-614-124B-803
; Sequence 803, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Panger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mennon, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 803
; LENGTH: 1238
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-803

Alignment Scores:
Pred. No.: 2,14e-37 Length: 1238
Score: 453.00 Matches: 120
Percent Similarity: 50.52% Conservative: 73

```

```

Best Local Similarity: 31.41% Mismatches: 143
Query Match: 13.25% Indels: 46
DB: 4 Gaps: 13

US-10-029-345A-109 (1-665) x US-09-614-124B-803 (1-1238)
Qy      19 LeuGluSerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsn 38
Db      168 CTCGCAAGCGGC---GGCAAGTCCCTGCTCTGATCTGACAGACCTTCTCGGCGCACAGC 224
Qy      39 ThrSerHisIleLeuGluAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeu 58
Db      225 GCGGCTACATCTTGAAGTTCGCTGCAAGTGCCTGTAAACCATCTGTCGCGCGCGG---281
Qy      59 GlnGlnAspLysValLeuIleThrGluLeuIle-----GlnHisSerAlaValHisLys 76
Db      282 GCTAAGGCTCCCTGAGCGCTGAGAGATCTGCCCCGCGAGAGAGAGTACCGCGCCGC 341
Qy      77 ValAspIleAspCysSerGlnLysValValIleTyrAspGlnSerSerGlnAspValAla 96
Db      342 TTCGCTCGGCGCTTACTCGCGGTATCTGTTACAGAGAGCGGCGCGCGCGAG 401
Qy      97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPhe---115
Db      402 AGCCTCGGAGAGACAGACCGGTGCTGCTGAGTGTGTCAGAGCGCTGCGCAAGCGCGAG 461
Qy      116 ---AsnSerValHisLeuLeuAlaGlyIlePheAlaGluPheSerArgCysPheProGly 134
Db      462 CGCACCGAATCTGCTGCTCAAGGCGGTAGAGAGGTTTCTTCGAGTACCGAGAA 521
Qy      135 LeuCysGluGlyLysSerThrLeu-----ValProThrCysIleSer 148
Db      522 TTCTGTTCTTAAACCAAGGCGCTGCGAGCATCCACCCCGGTTCGCCAGCGCACCA 581
Qy      149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db      582 GAGCCCTGGAGCTGAGTACAGCTCTGTTGGGAGCCCACTACAGACAGAGAGGCTCT 641
Qy      160 ThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuLeuLysGluLeu 179
Db      642 GTGAGATCTCTTCTCTTACTTACAGAGTGCCTTACCATCTGCCCCGAGAGACATG 701
Qy      180 IleGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAsp 199
Db      702 CTGAGAGCCCTGGGACTACAGGCTCTGTGATGTCTCTCGAGACTGCCAAAC---CAC 758
Qy      200 PheIIProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIle 219
Db      759 TTTGAAGACACTATCACTGACAGTGCATCCAGTGAAGATTAACCAAGCCGACATC 818
Qy      220 LeuProTrieLeuAspLysSerValAspPheIIleGluValAlaValAsnAngIyCys 239
Db      819 AGCTCTGTTGATGAGAGCCATAGAGTACATCCAGTGAAGAGACTCCGCGGCGC 878
Qy      240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIle 259
Db      879 GTGCTGTGCACTGCCAGGCGGCGCATCTCGCGGTCCGACCAACATCTCCCTGACTACTG 938
Qy      260 MetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIlyGluValArgPro 279
Db      939 ATGATGAAGAAACCGGAGGCTGAGGCTGAGAGAGGCTTCGAGTTCGTTAAGCAGCGCCGAGC 998
Qy      280 ThrIleSerProAsnPheAsnPheLeuGlyGluLeuAspTyrGluValIlyLys 299
Db      999 ATCATCTCGCCCAACTTCAGCTTCATGAGGAGCTGCTGAGTTCGAGTCCCAAGTCTG 1058
Qy      300 AsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysIleLeuGluLysPro 319
Db      1059 GCCACGCTCTGTGCTGCGGAGGCTGCTAGC-----CCC 1091
Qy      320 AsnGluProValProAlaValSerGluGlyGluLysSerGluThrProLeuSerPro 339
Db      1092 TCGGAGACCCCTG-----GGGAGCGGGGCAAGACCCCGCCAGCCACC 1133

```

```
QY 340 ProCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSer 359
Db 1134 -----ACCTGCAAGTTGCTTCAAGCTTTTCGGGTC---TCCGTGGGC 1172
QY 360 ValProSerValProSerValGlnProSerLeuLeuAlaAspSerProLeuValGlnAla 379
Db 1173 GTGCACCTCGGCCCCCAGACAGCCTGCTTACCTG-----CACAGCCCCATCACCACTCT 1226
QY 380 LeuSer 381
Db 1227 CCCAGC 1232

RESULT 10
US-09-671-325-803
; Sequence 803, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 803
; LENGTH: 1238
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-803

Alignment Scores:
Pred. No.: 2,14e-37 Length: 1238
Score: 453.00 Matches: 120
Percent Similarity: 50.52% Conservative: 73
Best Local Similarity: 31.41% Mismatches: 143
Query Match: 13.25% Indels: 46
DB: 4 Gaps: 13

US-10-029-345A-109 (1-665) x US-09-671-325-803 (1-1238)
QY 19 LeuGluSerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsn 38
Db 168 CTGCCGACGGG---GGCAAGTGGCTGCTGCTGACGCTGACAGCAGCTTCTCGCCACAGC 224
QY 39 ThrSerHisIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 58
Db 225 GCGGAGCTCAATCTGATGTTGCTGACGTCGCTGACGCTGACCAACATCGCGCGCGG--- 281
QY 59 GlnGlnAspLysValLeuIleThrGluLeuIle-----GlnHisSerAlaIleHis 76
Db 282 GCTAAGGGCTCCGAGGAGCTGAGAGATCTCTGCCCGCCAGAGAGAGATACGGCCCGC 341
QY 77 ValAspIleAspCysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAla 96
Db 342 TTGGGCTCGGCTCTACTCGCGCTCATCGCTTACGACGACGACGACGCGCGCGCGAG 401
QY 97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyIleGluLysSerPhe--- 115
Db 402 AGCTTCGCGGAGACAGACCGTGTGCTGAGGAGCGGCTGCGCCCAACGCGGAG 461
QY 116 ---AsnSerValHisLeuLeuAlaGlyIleGlyPheAlaGluPheSerArgCysPheProGly 134
Db 462 CGCAGCAGCATCTGCTGCTCAAAAGGCGGCTATGAGAGGTTTCTCCGAGTACCCAGAA 521
```

```
QY 135 LeuCysGluGlyLysSerThrLeu-----ValProThrCysIleSer 148
Db 522 TTTCGTTCTAAACCAAGGCTCTGGCAGCATCCACCCCGGTTCCCCAGGCGCAC 581
QY 149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db 582 GAGCCCTTGAACCTGACCTGAGCTCTGTGGAGACCCCACTACACAGACGAGAGGCTCT 641
QY 160 ThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGlyLeu 179
Db 642 GTGAGATCTCTTCCCTTCTCTACTGCGGAGTGGCTTACCATCATGCTGCGGAGACATG 701
QY 180 IleGlnIleAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAsp 199
Db 702 CTGAGCCCTCGGACATCACGGCTCTGTGTAATGTCTCTCGGACATGCCCAAC---CAC 758
QY 200 PheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 219
Db 759 TTTCAGAGACATATCATGATCAAGTGCATCCCATGGAAGATACCAAGCCGACATC 818
QY 220 LeuProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCys 239
Db 819 AGCTCCTGTTTCATGGAAGCCATAGATCATGATGCGGTGAAGGATGCGGTGGGCGC 878
QY 240 ValLeuValHisCysIleuAlaGlyIleSerArgSerHisThrIleAlaIleAlaTyrIle 259
Db 879 GTGCTGTGACATGCGACGCGGCAATCTGCGGTGCGCACCATCTGCTGGCTTACCTG 938
QY 260 MetLysArgMetAspSerLeuAspGluAlaTyrArgPheValLysGluLysArgPro 279
Db 939 ATGATGAAGAAACGGGTGAGGCTGAGAGAGGCTTGAATGCTTGAAGCAGCGCCGAC 998
QY 280 ThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLys 299
Db 999 ATCATCTCGCCCAACTCAGCTTCATGAGGCGGACGCTCTCAGTTCAGATGCCACAGTGTG 1058
QY 300 AsnGluIleValAspSerGlyProLysSerLysLeuLysLeuHisIleuGluLysPro 319
Db 1059 GCCACGCTCTGCTGCTCGGAGGCTGTAC-----CCG 1091
QY 320 AsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerPro 339
Db 1092 TCGGAGCCCTG-----GGGAGAGCGGAGCAAGACCCCGCACCCCGC 1133
QY 340 ProCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSer 359
Db 1134 -----ACCTGCAAGTTGCTTCAAGCTTTTCGGGTC---TCCGTGGGC 1172
QY 360 ValProSerValProSerValGlnProSerLeuLeuAlaAspSerProLeuValGlnAla 379
Db 1173 GTGCACCTCGGCCCCCAGACAGCCTGCTTACCTG-----CACAGCCCCATCACCACTCT 1226
QY 380 LeuSer 381
Db 1227 CCCAGC 1232

RESULT 11
US-09-589-184-803
; Sequence 803, Application US/09589184
; Patent No. 666447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C8
```

[illegible]

```

Db      939 ATGATGAAGAAACGGGTGAGCGCTTGAGAGAGCCCTTCAGTTCTGTAAGCAGCCGCCGACG  998
        |||          :|||          :|||          :|||          :|||          :|||
Qy      280 ThrlleserProkaphnePhleuglycInleuleumaptryrGIuylsyIlaelys  299
        |||          :|||          :|||          :|||          :|||          :|||
Db      999 ATCATCTCGCCCACTTGCACCTTCATTGGGCAAGCTGCTGAGATTCAGATCCCAAGTCTG  1058
        |||          :|||          :|||          :|||          :|||          :|||
Qy      300 AsnInthrGlyAlaserGlyProlysSerLyseuleuHlsleuGlulyePro  319
        |||          :|||          :|||          :|||          :|||          :|||
Db      1059 GCCACGTCCTGCTGCTGGGAGGCTGTAAC-----CCC  1091
        |||          :|||          :|||          :|||          :|||          :|||
Qy      320 AsnInlProValProAlaValaserGluGlyGlnLysSerGlnThrProLeuSerPro  339
        ::|          :|||          :|||          :|||          :|||          :|||
Db      1092 TCGGGAGACCCCTG-----GGGGAGCGGGGCAAGACCCTCCGACACCCCC  1133
        |||          :|||          :|||          :|||          :|||          :|||
Qy      340 ProCyalaabSeralatmrSerGluAlaIaelylnrGrProvalHisProalaser  359
        |||          :|||          :|||          :|||          :|||          :|||
Db      1134 -----ACCTCGCAGTTCGTCTTCACGCTTCCGGTCT--TCGGTGGGC  1172
        |||          :|||          :|||          :|||          :|||          :|||
Qy      360 ValProserValProserValGlnProSerLeuLeuGlnabSerProLeuValGlnAla  379
        |||          :|||          :|||          :|||          :|||          :|||
Db      1173 GTGACATCGGCCCCCAGCACGCTGCTTACTG-----CACAGCCCATCACAACCTCT  1226
        |||          :|||          :|||          :|||          :|||          :|||
Qy      380 LeuSer 381
        |||
Db      1227 CCCAGC 1232

RESULT 12
US-09-016-434-1100
; Sequence 1100, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Sellhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-4166
TELEFAX: (650) 845-0555
INFORMATION FOR SEQ ID NO: 1100:
SEQUENCE CHARACTERISTICS:
LENGTH: 2240 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1255784

```



```

Db      745 AGCTTCGGAGAGACGACCGCTGCTGTGTGTGAGAGCGGTGGCGGCAAGCGGAG 804
Qy      116 ---AasSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGly 134
Db      805 CGCAGCCGACATCTGCTGCTCAAAAGCGGCTATAGAGGTTTCTCCGAGTACCAGAA 864
Qy      135 LeuCyGluGlyLysSerThrLeu-----ValProThrCysIleSer 148
Db      865 TTCTGTTCTTAAACCAAGGCCCTGGAGCATCCACCCCGGTTCCTCCCACTGCCACA 924
Qy      149 GlnPro-----Cys---LeuProValAlaAsnIle---GlyPro 159
Db      925 GAGCCCTTGGACCTGGGCTGCACTCTGTTGGAGACCCCACTACACAGCAGGGGGTCT 984
Qy      160 ThrArgIleLeuProAsnLeuTyLeuGlyCysGlnArgAspValLeuAsnIleValLeu 179
Db      985 GTGGAGATCTCTTCTCTCTACCTCGAGTGCCTACCATCTGCTGCCGAGAGACATG 1044
Qy      180 ILeGlnAlaAsnGlyIleGlyTyValLeuAsnAlaSerTyThrCysProLysProAsp 199
Db      1045 CTGGAGCGCTTGGGACATCAGCGCTCTGTGAATGCTCTCGGACTGCCAAAC---CAC 1101
Qy      200 PheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 219
Db      1102 TTGTGAAGACACTATCATGTCAGTCAAGTCATCCCACTGGAGATTAACCAAGCCGACATC 1161
Qy      220 LeuProTyrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCys 239
Db      1162 AGCTCTGTGTTCTATGAGACCATAGATCATGATGATGCGTGAAGACTGCCGTGGGCGC 1221
Qy      240 ValLeuValHisCysLeuAlaGlyIleSerArgSerPheAlaThrIleAlaIleAlaTyIle 259
Db      1222 GTCTGTGTGACTGACCGAGCGGAGATCTCGCGGTCCGACCAATCTGCTGCGCTTACTG 1281
Qy      260 MetLysArgMetAspSerLeuAspGluAlaTyThrArgPheValLysGluLysArgPro 279
Db      1282 ATGATATGAAGAAAGGGGTGAGAGCTGGAGAGCCCTTGAGTTCTGTAAAGCAGCCCGCAGC 1341
Qy      280 ThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyGluLysIleLys 299
Db      1342 ATCATCTCGCCCAACTTCACTTCAATGGGAGCACTGCTGATGATGCTCCAGTCCAGTGC 1401
Qy      300 AsnGlnThrGlyLysSerGlyProLysSerLysLeuLysLeuLysIleLysPro 319
Db      1402 GCCACGTCTCTGCTGCGAGGCTGCTAGC-----CCC 1434
Qy      320 AsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSer-Pr 339
Db      1435 TCGGGACCCCTG-----CGGAGCGGGGCAAGACCCCGCCACCCCC 1476
Qy      339 oProCyAlaAspSerAlaThr-----SerGluAlaAlaGlyGlnArgProValHisPr 357
Db      1477 ACCTTCAGATGTCCTTCACTTCCGAGTCTCCGATGGGCGTGCACTCGGCC-----CC 1530
Qy      357 oAlaSerValProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuVa 377
Db      1531 AGCAGCTGCTTCACTGACGACAGCCCATCAACCACTCTC-----CCAGCTGT 1578
Qy      377 IGlAlaLeuSerGly 382
Db      1579 TAGAGCCGCCCTGGGG 1594

```

```

; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Ligu
; APPLICANT: Wang, Aljun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 825
; LENGTH: 2064
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-736-457-825

Alignment Scores:
Pred. No.: 1 02e-36 Length: 2064
Score: 450.50 Matches: 122
Percent Similarity: 51.81% Conservative: 78
Best Local Similarity: 31.61% Mismatches: 140
Query Match: 13.18% Indels: 46
DB: 4 Gaps: 14

US-10-029-345A-109 (1-665) x US-09-736-457-825 (1-2064)
Qy      19 LeuGluSerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyAsn 38
Db      511 CTGCGAGCGGC---GGCAAGTGCCTGCTGTGATGTCCTCGGACTGCCAGACGTTCTCGGCGCAGC 567
Qy      39 ThrSerHisIleGluGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgGluLeu 58
Db      568 GCGGGCTATCTTCAATGTTGCTGTCACGTCGCTGTAAACCAATCTGCGCGCGCGG--- 624
Qy      59 GlnGlnAspLysValLeuIleThrGluLeuIle-----GlnHisSerAlaLysHisLys 76
Db      625 GCTAAGGCTCCGTGAGCTTGAAGCATCTTCCCGCGAGAGAGAGAGATCCGCCCGC 664
Qy      77 ValAspIleAspCysSerGlnLysValValAlaTyThrAspGlnSerGlnAspValAla 96
Db      685 TTGCGCTCGGCGCTTACTCGGCGGTCACTGCTCACTGACAGAGCGGAGCGCGCGCAG 744
Qy      97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPhe--- 115
Db      745 AGCTTCGGAGAGACGACCGGTGCTGTGTGTGTCAGAGCGCTGCGCGCAAGCGGAG 804
Qy      116 ---AasSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGly 134
Db      805 CGCAGCCGACATCTGCTGCTCAAAAGCGGCTATAGAGGTTTCTCCGAGTACCAGAA 864
Qy      135 LeuCyGluGlyLysSerThrLeu-----ValProThrCysIleSer 148
Db      865 TTCTGTTCTTAAACCAAGGCCCTGGAGCATCCACCCCGGTTCCTCCCACTGCCACA 924
Qy      149 GlnPro-----Cys---LeuProValAlaAsnIle---GlyPro 159
Db      925 GAGCCCTTGGACCTGGGCTGCACTCTGTTGGAGACCCCACTACACAGCAGGGGGTCT 984
Qy      160 ThrArgIleLeuProAsnLeuTyLeuGlyCysGlnArgAspValLeuAsnIleValLeu 179
Db      985 GTGGAGATCTCTTCTCTCTACCTCGAGTGCCTACCATCTGCTGCCGAGAGACATG 1044
Qy      180 ILeGlnAlaAsnGlyIleGlyTyValLeuAsnAlaSerTyThrCysProLysProAsp 199
Db      1045 CTGGAGCGCTTGGGACATCAGCGCTCTGTGAATGCTCTCGGACTGCCAAAC---CAC 1101
Qy      200 PheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 219
Db      1102 TTGTGAAGACACTATCATGTCAGTCAAGTCATCCCACTGGAGATTAACCAAGCCGACATC 1161
Qy      220 LeuProTyrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCys 239

```


Qy 377 |G|n|A|Leu|ser|G|Y 382
|:::|
Db 1579 TAGAGCCGCCCTGGGG 1594

Search completed: June 22, 2004, 02:46:14
Job time : 150.037 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:13:37 ; Search time 14.4416 Seconds

(without alignments)
2397.707 Million cell updates/sec

Title: US-10-029-345A-109

Sequence: 1 MAHEMIGTQIVTERVALLE.....LGKVGSSSPSSGMEIIEVVS 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3406	99.6	665	1 DUSG_HUMAN	Q9b9g4 homo sapien
2	1326	38.8	625	1 DUS8_HUMAN	Q13202 homo sapien
3	1302	38.1	663	1 DUS8_MOUSE	009112 mus musculi
4	497.5	14.6	619	1 VHP1_CAEEL	010038 caenorhabd
5	471.5	13.8	381	1 DUS6_HUMAN	Q16828 homo sapien
6	469	13.7	483	1 DUS6_MOUSE	Q9b66 mus musculi
7	468	13.7	483	1 DUS6_MOUSE	Q9b66 mus musculi
8	466.5	13.6	381	1 DUS6_MOUSE	Q9b66 mus musculi
9	465.5	13.6	381	1 DUS6_MOUSE	Q9b66 mus musculi
10	445	13.0	394	1 DUS4_HUMAN	Q13115 homo sapien
11	444	13.0	375	1 DUS4_HUMAN	Q9b71 gallus gall
12	439.5	12.9	395	1 DUS4_HUMAN	Q9b71 gallus gall
13	433	12.7	367	1 DUS1_MOUSE	P28563 mus musculi
14	423	12.4	367	1 DUS1_MOUSE	P28563 mus musculi
15	420	12.3	367	1 DUS1_MOUSE	P28563 mus musculi
16	418	12.2	367	1 DUS1_MOUSE	P28563 mus musculi
17	417	12.2	367	1 DUS1_MOUSE	P28563 mus musculi
18	409	12.0	318	1 DUS2_HUMAN	Q05922 mus musculi
19	390	11.4	320	1 DUS7_HUMAN	Q16829 mus musculi
20	386.5	11.3	320	1 DUS7_MOUSE	Q1426 mus musculi
21	367	10.7	384	1 DUS5_MOUSE	Q54838 mus musculi
22	361.5	10.6	280	1 DUS7_MOUSE	Q54838 mus musculi
23	360.5	10.5	280	1 DUS7_MOUSE	Q54838 mus musculi
24	269	7.9	198	1 DUS5_HUMAN	Q16650 mus musculi
25	268	7.8	198	1 DUS5_HUMAN	Q16650 mus musculi
26	205.5	6.0	188	1 DUS1_MOUSE	Q91147 mus musculi
27	205.5	6.0	217	1 DUS1_MOUSE	Q91147 mus musculi
28	205.5	6.0	223	1 STYX_HUMAN	Q8W12 homo sapien
29	205	6.0	223	1 STYX_HUMAN	Q8W12 homo sapien
30	204	6.0	278	1 FMPI_MOUSE	Q80969 mus musculi
31	202	5.9	364	1 FMPI_MOUSE	Q80969 mus musculi
32	201.5	5.9	313	1 FMPI_MOUSE	Q80969 mus musculi
33	196	5.7	185	1 DUS3_HUMAN	Q9d7x3 mus musculi

34	194.5	5.7	185	1 DUS3_MOUSE	Q9d7x3 mus musculi
35	189	5.5	489	1 MS5_YEAST	P3850 saccharomyc
36	185	5.4	807	1 YBPT_YEAST	P38148 saccharomyc
37	183.5	5.4	276	1 PTP3_CHIEU	Q39431 chlamydomon
38	181.5	5.3	295	1 DUSF_HUMAN	Q9h12 homo sapien
39	177	5.2	125	1 DUSF_MOUSE	Q8442 mus musculi
40	171.5	5.0	339	1 DUSC_MOUSE	Q9d02 mus musculi
41	170.5	5.0	292	1 Y042_CAEEL	P34680 caenorhabd
42	161	4.7	198	1 DUSC_HUMAN	Q9u16 homo sapien
43	158.5	4.6	340	1 DUSC_HUMAN	Q9u16 homo sapien
44	157	4.6	209	1 Y113_YEAST	P40479 saccharomyc
45	156	4.6	198	1 DUSD_MOUSE	Q9y17 mus musculi

ALIGNMENTS

RESULT 1
DUSG_HUMAN STANDARD; PRT; 665 AA.

AC Q9BY64; Q9COG3; 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase
DE phosphatase 7) (MKP-7).
GN DUSP16 OR MKP7 OR KIAA1700.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,
RT functions as a shuttle protein";
RT J. Biol. Chem. 276:39002-39011(2001). October
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21082932; PubMed=11214970;
RA Niggate T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:347-355(2000).
CC -|- FUNCTION: Involved in the inactivation of MAP kinases.
CC -|- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -|- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -|- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -|- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AB052156; BAB40814.1; -;
CC EMBL: AB051487; BAB21791.1; ALT_INIT.
CC HSSP: Q16828; IMKP.
CC Genew: HGNC:17909; DUSP16.
CC MIM: 607175; -;
CC GO: GO:0005737; C:cytoplasm; TAS.
CC GO: GO:0005634; C:nucleus; TAS.

DR GO:0004721; F:protein phosphatase activity; TAS.
 DR GO:0016311; P:dephosphorylation; TAS.
 DR GO:0000188; P:inactivation of MAPK; TAS.
 DR GO:0045209; P:leptomycin B-sensitive MAPK phosphatase nuc. . .; TAS.
 DR GO:0045204; F:MAPK nucleus export; TAS.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00195; DSPC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS50206; RHODANES_3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolase; Nuclear protein.
 FT DOMAIN 22 137 RHODANES.
 FT DOMAIN 227 289 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 244 244 PHOSPHOCYSTEINE INTERMEDIATE (BY SIMILARITY)

SEQUENCE 665 AA; 73101 MW; 1BD853FF08460DF CRC64;
 Query Match 99.6%; Score 3406; DB 1; Length 665;
 Best Local Similarity 99.7%; Pred. No. 7, 7e-200;
 Matches 665; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALESGTEKLLIDSRPVEYNTSHLEAININCSGLMRRIQQ 60
 DB 1 MAHEMIGTQVTERLVALESGTEKLLIDSRPVEYNTSHLEAININCSGLMRRIQQ 60
 QY 61 DKLTITELIOHSAKHKVIDCSQKVYVQSSQDVASISDCEFLVLLGKLEKSNVHL 120
 DB 61 DKLTITELIOHSAKHKVIDCSQKVYVQSSQDVASISDCEFLVLLGKLEKSNVHL 120
 QY 121 LAGGFAPSRCPGCEGKSTVPTICISPCIPVANIPTLLENLYGCGRDVANKELI 180
 DB 121 LAGGFAPSRCPGCEGKSTVPTICISPCIPVANIPTLLENLYGCGRDVANKELI 180
 QY 121 LAGGFAPSRCPGCEGKSTVPTICISPCIPVANIPTLLENLYGCGRDVANKELI 180
 DB 121 LAGGFAPSRCPGCEGKSTVPTICISPCIPVANIPTLLENLYGCGRDVANKELI 180
 QY 181 QONGIGYVLAASVTPCKPDPFIPESHFLRPVNDSCCEKILPMWLSKVDPIEKARANGCV 240
 DB 181 QONGIGYVLAASVTPCKPDPFIPESHFLRPVNDSCCEKILPMWLSKVDPIEKARANGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTTSPNFELGQLLDYEKKIKN 300
 DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTTSPNFELGQLLDYEKKIKN 300
 QY 301 QTGAGGPKSKLKLHLKEKNEPVPAVSEGGQKSTPLSPCADSATSEAGQRPVHPASV 360
 DB 301 QTGAGGPKSKLKLHLKEKNEPVPAVSEGGQKSTPLSPCADSATSEAGQRPVHPASV 360
 QY 361 PSVPSVQSLLEDSPVQALSGHLISADRLSDSNKLKRSFSLDIKSYVSASMAASLHGF 420
 DB 361 PSVPSVQSLLEDSPVQALSGHLISADRLSDSNKLKRSFSLDIKSYVSASMAASLHGF 420
 QY 421 SSSSDALEYKPSFTLLDGTNKLCOFSPVOELSBOTPTSTDKERASIPKKLQTPARPSDQ 480
 DB 421 SSSSDALEYKPSFTLLDGTNKLCOFSPVOELSBOTPTSTDKERASIPKKLQTPARPSDQ 480
 QY 481 SKRLHVSRTSSSGTAQSSLLSPLHRSQSVEDNYHTSFLGLSTSQOHLTKSAGLGLKGMH 540
 DB 481 SKRLHVSRTSSSGTAQSSLLSPLHRSQSVEDNYHTSFLGLSTSQOHLTKSAGLGLKGMH 540
 QY 541 SDLIAPOTSTPLSSVYFATSESHPSASAIYGSASISYSCGLPTCGQVYSVRRR 600
 DB 541 SDLIAPOTSTPLSSVYFATSESHPSASAIYGSASISYSCGLPTCGQVYSVRRR 600
 QY 601 QKPSDRADSRSSWHESSPFEKQPKRRSCOMEFGESIMSENRSEBELGKVGSSSPSSGME 660
 DB 601 QKPSDRADSRSSWHESSPFEKQPKRRSCOMEFGESIMSENRSEBELGKVGSSSPSSGME 660

QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 2
 DUS8_HUMAN STANDARD; PRT; 625 AA.
 ID DUS8_HUMAN
 AC Q13202;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Dual specificity protein phosphatase hvh-5).
 GN DUSP8 OR VH5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=96009533; PubMed=7561881;
 RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.;
 RT "hvh-5: a protein tyrosine phosphatase abundant in brain that
 RT inactivates mitogen-activated protein kinase.";
 RL J. Neurochem. 65:1823-1833(1995).
 CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
 CC phosphate as well as with serine/threonine-protein phosphatase (by
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- TISSUE SPECIFICITY: Abundant in brain, heart and skeletal muscle.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U27193; AA83151.1; -;
 DR HSSP: Q16828; IMRP.
 DR Genew; HGNC:3074; DUSP8.
 DR MIM: 602038; -;
 DR GO: GO:0005737; C:cytoplasm; TAS.
 DR GO: GO:0004725; F:protein tyrosine phosphatase activity; TAS.
 DR GO: GO:0000188; P:inactivation of MAPK; TAS.
 DR GO: GO:0006470; P:protein amino acid dephosphorylation; TAS.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00195; DSPC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS50206; RHODANES_3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolase; Nuclear protein.
 FT DOMAIN 23 138 RHODANES.
 FT DOMAIN 162 430 PROTEIN-TYROSINE PHOSPHATASE.
 FT DOMAIN 310 550 PRO-RICH.
 FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY SIMILARITY).

SQ SEQUENCE 625 AA; 65840 MW; DOBEAL44487219666 CRC64;
 Query Match 38.8%; Score 1326; DB 1; Length 625;
 Best Local Similarity 45.1%; Pred. No. 2,4e-73;
 Matches 314; Conservative 94; Mismatches 185; Indels 104; Gaps 21;

QY 1 MAHEMIGTOIV-TERIVALLSEGEKVLIDSRPFVYNTSHILEAININCSKLMKRRLO 59
 1 MADDRIPRKVMADKLASLLRGPGPLVIDSRSPFVYNSCHVLSVNICSLTVKRRLO 60
 DB 1 MADDRIPRKVMADKLASLLRGPGPLVIDSRSPFVYNSCHVLSVNICSLTVKRRLO 60
 QY 60 QDRVLTTELIOHSAKHVIDSCQKVVYVYDQSDVASLSSDCEFLTVLGLKLEKSPSVH 119
 61 QGVTTIAELIQPARSGVEATEPQDVVYDQSTRDASVLAADSPFLSLKLDGCFDSVA 120
 QY 120 LLAGFAEBSRCPPGLCEGK-STLVPTICISQPLCPVANITPTRLPLNLYGCGQDVLANKE 178
 121 ILTGFAFTSSCPGGLCEGKPAALPMSLSQPCLPVPSVGLTILPLTLVLSGSKDVLNKO 180
 QY 179 LIOONGIGVYLVNAYTCRKPDPFIPESHFLRPVNDSCFCEILPWLKSDVFIRKASNG 238
 181 LMTQNGISYLVNANSNCPKDPFICBSRFMRVPINDNCEKLLPWLKSLFIDKALSSC 240
 QY 239 CVLVHCLAGISRATTAIAIYIMKMDSLDEAYRVEKRPISPNFPLGOLLDEYEKKI 298
 241 QVIVHCLAGISRATTAIAIYIMKMDSDAVRFVDRRPSISPNFPLGOLLDEYERL 300
 QY 299 KNOTGASGPKSKLKLHLEKPN--EPVAVSBCGQKSEPLPSPCADSAT--SEAAQR 353
 301 KLALALQGDG-----TPSGTPEPPPPAPAGAPLRPLP--PPTSESATGNAAAREG 350
 QY 354 PVHPSVSPSVQPSLLEDSPLVQALSGHLSDRLDENKLLKRSPLIKSVYSASM 413
 351 GLSAGGPPAPPPPPA--TSALQQGRGLHLSDRLODTNRLKRSPLIKSA----- 401
 QY 414 AASLHGSSSEDALEYYKPESTLDTGN-----KLQCF--SP--VQLESEOTPEPSPD 461
 402 -----YAPRRRDPGPPPPPPGPAKCKLSDSGALGSSSPD--SPD 444
 QY 462 KEBASIPKQLQTPAPSDSQKRLHVSRTSSGTRQRLSLPLHSGSVEDNHTSPLFG- 520
 445 AAPEARPRRRPRP-----PAGSPAR--SPASILG-----LNFCD 478
 QY 521 --LSTSQHLLTKAGLGLG-----WHDILAQSTPSLTSWYFATESSHRYSA 569
 479 AARQTPPHGISALSAFGLPPGQAPGAWAPLDSF--GTPSPDGCWCFSP----- 529
 QY 570 SAIVGSGASVAYSCQLPTCGDQVYVVRROKPSDRADRSRWHEESPPEKQPKRRSCQ 629
 530 GAGAGAGVLFAPGRACAPGPGGSDLRREBAARAEPRDARTGMPPEPAPETQPKRRSCQ 589
 QY 630 MEFGESIMSENRR--BELGKVGSSQSPSGSMETIEVS 665
 590 MEFEPEG-MVEGRARGELALGKQSPSGSVEVEVS 625
 DB

RESULT 3
 DUS8_MOUSE STANDARD; PRT; 663 AA.
 AC 009112;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
 GN (Neuronal tyrosine threonine phosphatase 1).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteleota; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96311565; PubMed=8733137;

RA Theodosiou A.M., Rodrigues N.R., Nesbit M.A., Ambrose H.J.,
 RA Paterson H., McEltan Arnold E., Boyd Y., Levertha M.A., Owen N.,
 RA Blake D.J., Ashworth A., Davies K.E.,
 RT "A member of the MAP kinase phosphatase gene family in mouse
 RT containing a complex trinucleotide repeat in the coding region";
 RL Hum. Mol. Genet. 5:675-684 (1996).
 CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
 CC phosphate as well as with serine/threonine-protein phosphatase (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
 CC -1- TISSUE SPECIFICITY: Expressed predominantly in brain and lung.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X95518; CAA64772.1; -.
 CC HSSP: Q16828; IMKP.
 CC MGD: MGI:106626; Dusp8.
 CC InterPro: IPR000340; DS_phosphatase.
 CC InterPro: IPR008343; MAPK_phosph.
 CC InterPro: IPR001763; Rhodanese-like.
 CC InterPro: IPR000387; TYR_phosphatase.
 CC Pfam: PF00782; DSPC; 1.
 CC PRINTS: PR01764; MAPKPHPTASE.
 CC SMART: SM00195; DSPC; 1.
 CC SMART: SM00450; RHOD; 1.
 CC PROSITE: PS50206; RHODANES_3; 1.
 CC PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 CC PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 CC PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 CC KMW Hydrolyase; Nuclear protein.
 CC FT DOMAIN 23 138 RHODANES.
 CC FT DOMAIN 162 432 PROTEIN-TYROSINE PHOSPHATASE.
 CC FT DOMAIN 452 459 POLY-ARG.
 CC FT DOMAIN 555 558 POLY-SER.
 CC FT DOMAIN 576 576 POLY-GLY.
 CC FT DOMAIN 577 600 POLY-SER.
 CC FT DOMAIN 311 552 PRO-RICH.
 CC FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
 CC SIMILARITY).
 SQ SEQUENCE 663 AA; 68847 MW; 416P429A12C1FA7C CRC64;
 Query Match 38.1%; Score 1302; DB 1; Length 663;
 Best Local Similarity 43.4%; Pred. No. 7,4e-72;
 Matches 318; Conservative 102; Mismatches 176; Indels 136; Gaps 26;

QY 1 MAHEMIGTOIV-TERIVALLSEGEKVLIDSRPFVYNTSHILEAININCSKLMKRRLO 59
 1 MADDRIPRKVMADKLASLLRGPGPLVIDSRSPFVYNSCHVLSVNICSLTVKRRLO 60
 DB 1 MADDRIPRKVMADKLASLLRGPGPLVIDSRSPFVYNSCHVLSVNICSLTVKRRLO 60
 QY 60 QDRVLTTELIOHSAKHVIDSCQKVVYVYDQSDVASLSSDCEFLTVLGLKLEKSPSVH 119
 61 QGVTTIAELIQPARSGVEATEPQDVVYDQSTRDASVLAADSPFLSLKLDGCFDSVA 120
 QY 120 LLAGFAEBSRCPPGLCEGKSTLVPT-CTSQPLCPVANITPTRLPLNLYGCGQDVLANKE 178
 121 ILTGFAFTSSCPGGLCEGKPAALPMSLSQPCLPVPSVGLTILPLTLVLSGSKDVLNKO 180
 QY 179 LIOONGIGVYLVNAYTCRKPDPFIPESHFLRPVNDSCFCEILPWLKSDVFIRKASNG 238
 181 LMTQNGISYLVNANSNCPKDPFICBSRFMRVPINDNCEKLLPWLKSLFIDKALSSC 240
 DB

```

QY 239 CVLVHCLAGISRSATIAIAYIMKMDMSLDEAFYKREKPTISPNFNFLGQLLYEKKI 298
D 241 QVIVHCLAGISRSATIAIAYIMKMDMSLDEAFYKREKPTISPNFNFLGQLLYEKKI 300
QY 299 KNGTG--ASGPCKKLLHLEKNEBPVAVSEGGQKSEPTLPPCADSAT--SEA----- 350
D 301 KLAALQDTDP-----HLGTP--EPLMGPAAGIPLRPP--PSSSESAATGSEATATAR 350
QY 351 -----GQPVPAVSPVSPVQSLLEDSPVQALSGHLASADRLSDSKLRPSFLD 403
D 351 EGSPSAGGAPF--PSTAPATSLQ-----OGLGHLHSSDRLQDTYRLKRSFLD 399
QY 404 IKSIVTSASMAASLHGSSSEDALEYKPTLDTGNLQCFSPVQSLSEQTPETSDE 463
D 400 IKS-AVAPSRPPDPFPPDPGAPKLCGLDPSGGTGL-----PSPSPSP 445
QY 464 EASIPK---KLTARPSPDSQSKR--LHSAVTSISSTQAKRLSLPLHSGSVEDNYHTSFL 518
D 446 D-SVPECRPRRRRRPPASPPASPAHGLGFGDTAAQ--TPRH----- 487
QY 519 FGLSTSQOHLTKSAGL-----GLKGMHSDILAPOTSTPSTSWYFATESH----- 567
D 488 -GLSA-----LSAPFLPQPQAPGAPGCVVPLDSP--CTSPDGPWCFSBGAQGPQAVF 539
QY 568 SA-----SAYYG-----SASYSAYSCSGLPTCGDQV 594
D 540 SAFGRVSGAPGPNSSSSGGGGGGGGGGGGGGSSSSSSSSSSSSSSSSSSSSSSSS 599
QY 595 YSVRRQKPSDRASSRSMHESPEFQKFRKSCOMFEGSINSENR--RELGKKGSGS 653
D 600 SDRRR-----DVRGTGPEEPADAOFRKRCQHFEBG--WVGRAGGEELALGKQT 651
QY 654 SFGSGMEIIEVS 665
D 652 SFGSGVEIIEVS 663

```

RESULT 4

VHPI_CABEL

ID VHPI_CABEL STANDARD; PRT; 619 AA.

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein-tyrosine phosphatase vhp-1 (EC 3.1.3.48).
GN vhp-1 OR F08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Polodetidae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]_TaxId=6239;
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RA Chisoe S.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -!- SIMILARITY: belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23178; AAC46719.1; -
DR PIR; T15969; T15969.
DR HSSP; O16828; 1MKP.
DR WormPep; F08B1.1; CE01899.

```

```

DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00506; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00504; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase.
FT DOMAIN 200 269 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 92 95 POLY-SER.
FT DOMAIN 351 354 POLY-SER.
FT DOMAIN 465 472 POLY-SER.
FT DOMAIN 483 488 POLY-SER.
FT ACT_SITE 224 224 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 619 AA; 66354 MW; 369E326F615D0529 CRC64;

```

Query Match 14.6%; Score 497.5; DB 1; Length 619;
Best Local Similarity 28.9%; Pred. No. 5,8e-23;
Matches 160; Conservative 77; Mismatches 159; Indels 157; Gaps 19;

```

QY 123 GGFARFCRCPGLCEGKS--TLVPTCIQPC--PVANIGPRLIPNLYLGGQDVLNKE 178
D 99 GGFQFAQYFQLCESSEKMTLPGLSQPCLSQPTGD--GITLIPNLYLGGQDVLNKE 157
QY 179 LIQONGIGVYLVNASTCPKRPDIPE--SHFLRVVNDSPCEKILPMLDSDVPDIKAKSN 237
D 158 MLDALDISVIVINISMCKPSKVCIKEDKNFMRIPVNDVQEKLSFFPMAYIELEKCRAG 217
QY 238 GCVLVHCLAGISRSATIAIAYIMKMDMSLDEAFYKREKPTISPNFNFLGQLLYEKKI 296
D 218 KCLHCLAGISRSATIAIAYIMKMDMSLDEAFYKREKPTISPNFNFLGQLLYEKKI 277
QY 297 ---KIKNOTGASGPKSKLKL--HLEKNEP-----VPANSEG 329
D 278 LKIDHLDYDQASRPHRDYDGPDDLCPKPKASASCNCPGSTHDESSPSVSEG 337
QY 330 GQKSEPTLSPCADATSEAGQRPVPAVSPVQSLLEDSPVQALSGHLHSDRL 389
D 338 SAASEPTLSSAASSSTASA-----PPSMPTSTSE-----QGTSSGTVVNG 379
QY 390 LEDSNKLRPSFLDI-----KSVYSASMAASLHGSSSEDALEYKPTLDTGNLKC 443
D 380 -----KRNMTMDLGPDRKALGLESRICTSV-----ALPSPSTELSRLS--- 420
QY 444 QFSPVQSLSEQTP-----EISPDKEASIPKQLQTPAPSDQSKRLHSAVTSISSTQAKRSL 500
D 421 -FNGPEALAPSTPIILNFTNPFNSPIIPV-----ASSSREVLITLP 460
QY 501 SPLHSGSVEDNYHTSFLGISTSQOHLTKSAGLGLKGMHSDILAPOTSTPSTSWYF 560
D 461 TPAASSSS-----STS-----SEPSDFSSPES 483
QY 561 TESH-----FYSAGIYGSASYSAYSCGLPTCGDQVSVRRQKPSDADRRSMH 614
D 484 SSSSIYVENPFASITVPASSSISTPSGSG-----STPAASSSSAARSC 529
QY 615 EESPFQKFRKS 627
D 530 RMKGFFVFSKKA 542

```

RESULT 5

ID DUS6_HUMAN

AC Q16828; 075109; 09BSH6; STANDARD; PRT; 381 AA.

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
DE phosphatase 3) (MKP-3) (Dual specificity protein phosphatase PYST1).
GN DUSP6 OR MKP3 OR PYST1.

```


OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Forebrain;
 RX MEDLINE=96312959; PubMed=8670865;
 RA Groom L.A., Sneddon A.A., Alessi D.R., Dowd S., Keyse S.M.;
 RT "Differential regulation of the MAP, SAP and RK/p38 kinases by Pyk1,
 RT a novel cytosolic dual-specificity phosphatase.";
 RL EMBO J. 15:3621-3632(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Liver;
 RX MEDLINE=99077745; PubMed=9858808;
 RA Furukawa T., Yatsunaka T., Yousef E.M., Abe T., Yokoyama T.,
 RA Fukuhige S., Soeda E., Hoshi M., Hayashi Y., Sunamura M., Kobari M.,
 RA Horii A.;
 RT "Genomic analysis of DUSP6, a dual specificity MAP kinase phosphatase,
 RT in pancreatic cancer.";
 RL Cytogenet. Cell Genet. 82:156-159(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT VAL-114.
 RC TISSUE=Colon, Kidney, Skin, and Stomach;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Udell T.B., Tothiyuki S., Carinci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Halton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzyzanski M.I., Skalek U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 204-347.
 RX MEDLINE=99140299; PubMed=10048930;
 RA Stewart A.E., Dowd S., Keyse S.M., McDonald N.Q.;
 RT "Crystal structure of the MAPK phosphatase Pyk1 catalytic domain and
 RT implications for regulated activation.";
 RL Nat. Struct. Biol. 6:174-181(1999).
 CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
 CC family.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q16828-1; Sequence=Displayed;
 CC Name=2; Synonyms=DUSP6-ALT;
 CC IsoId=Q16828-2; Sequence=VSP_005137;
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; X93920; CAA63813.1; -;
 CC DR EMBL; AB013601; BAA31968.1; -;
 CC DR EMBL; AB013603; BAA31968.1; JOINED.
 CC DR EMBL; AB013600; BAA31968.1; JOINED.
 CC DR EMBL; AB013582; BAA31969.1; -;
 CC DR EMBL; AB013602; BAA31969.1; -;
 CC DR EMBL; BC003143; AAH03143.1; -;
 CC DR EMBL; BC003562; AAH03562.1; -;
 CC DR EMBL; BC005047; AAH05047.1; -;
 CC DR EMBL; BC037236; AAH37236.1; -;
 CC DR EMBL; IMKP; 22-JUL-99.
 CC DR PDB; 1H2M; 25-JAN-02.
 CC DR Genew; HGNC:3072; DUSP6.
 CC DR MIM; 602748; -;
 CC DR GO; GO:0005737; Cytoplasm; IDA.
 CC DR GO; GO:0005625; Cytosol fraction; TAS.
 CC DR GO; GO:0004722; F-protein serine/threonine phosphatase activity; IDA.
 CC DR GO; GO:0004725; F-protein tyrosine phosphatase activity; IDA.
 CC DR GO; GO:0000188; P-inactivation of MAPK; IDA.
 CC DR GO; GO:0006470; P-protein amino acid dephosphorylation; TAS.
 CC DR InterPro; IPR000340; DS phosphatase.
 CC DR InterPro; IPR008343; MAPK_phosph.
 CC DR InterPro; IPR001763; Rhodanese-like.
 CC DR InterPro; IPR000387; TYR_phosphatase.
 CC DR Pfam; PF00782; DSPC; 1.
 CC DR Pfam; PF00581; Rhodanese; 1.
 CC DR PRINTS; PRO1764; MARKPHRTASE.
 CC DR SMART; SM00195; DSPC; 1.
 CC DR SMART; SM00450; RHOD; 1.
 CC DR PROSITE; PS50206; RHODANSE_3; 1.
 CC DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
 CC DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 CC DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 CC KW Hydrolase; Polymorphism; Alternative splicing; 3D-structure.
 CC FT DOMAIN 30 148
 CC FT RHODANSE.
 CC FT PROTEIN-TYROSINE PHOSPHATASE.
 CC FT PROTEIN-TYROSINE PHOSPHATASE.
 CC FT ACT SITE 293 293
 CC FT VARSPLIC 134 279
 CC FT Missing (in isoform 2).
 CC FT /FTID=VSP_005137.
 CC FT L->V.
 CC FT /FTID=VAR_015113.
 CC FT VARIANT 114 114
 CC FT STRAND 208 211
 CC FT TURN 212 213
 CC FT STRAND 214 217
 CC FT TURN 219 220
 CC FT HELIX 225 230
 CC FT TURN 231 232
 CC FT STRAND 233 238
 CC FT STRAND 246 250
 CC FT TURN 251 252
 CC FT STRAND 253 257
 CC FT TURN 254 265
 CC FT HELIX 269 271
 CC FT HELIX 272 284
 CC FT TURN 285 286
 CC FT STRAND 288 292
 CC FT STRAND 298 312
 CC FT HELIX 298 312
 CC FT TURN 313 315
 CC FT HELIX 316 326
 CC FT TURN 328 329
 CC FT TURN 335 336
 CC FT HELIX 337 345
 CC FT TURN 346 346
 CC SQ SEQUENCE 381 AA; 42333 MW; 03BC12252CE73B26 CRC64;
 CC Query Match 13.8%; Score 471.5; DB 1; Length 381;
 CC Best Local Similarity 34.0%; Pred. No. 1,1e-21;
 CC Matches 113; Conservative 63; Mismatches 107; Indels 49; Gaps 8;

```

QY 19 LBSGTEKVLIDSPFVEYNTSHILEAININCSKLMKRLQODKVTITELIOHSA-KHKY 77
DB 27 LEIGMERILLMDCRQGLYESHSHESAINVAIPGIMRLRQKGLVPRAALFTGRDRF 86
QY 78 DIDC-SQKVVYVDSQSO-VASLSSDCLTVLLGKLEKSPNSVHLLAGGFAEFSRCFGL 135
DB 87 TRRCITDVVLYDESSDMNENTGESILGLLKKLMDCECRAFLYEGGFSKQAEFSLH 146
QY 136 CEKSTLVPYC-ISPCLPVANIG----- 158
DB 147 CE--TNLDGSCSSSPPLVGLGGLRISSDSSDIEDIDRPNATDSDSPSLSNQP 204
QY 159 --PTRIPLNYLGGQGVVANKELIQNGIGVYLNASVYTCPR-DFIPESHFLRPVUNSF 215
DB 205 SFPFELLIPVLYGCAKSTNLDVLEERGIKYLIVLPNLPLFNAGEFKQKQIPISDH 264
QY 216 CEKLPVLDSVDFIEKAKASNCVHCLAGISRSATIAIYMKRDMGLDEAYRFRK 275
DB 265 SONLSQFPFPAISFIDBARGKCGVAVHCLAGISSVTVVAVYLMQKLMSMDAYDIVK 324
QY 276 EKPTTSPNENFGQLIDVEKIKRGTGASGP 307
DB 325 MKKSNISPNENFGQLIDERTL----GLSSP 352

```

RESULT 6

```

DUSA_HUMAN STANDARD; PRT; 482 AA.
AC G9Y6W6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase
DE phosphatase 5) (MKP-5).
GN DUSP10 OR MKP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE=99321929; PubMed=10391943;
RA Tanoue T., Moriyuchi T., Nishida E.;
RT "Molecular cloning and characterization of a novel dual specificity
RT phosphatase, MKP-5."
RL J. Biol. Chem. 274:119949-119956(1999).
RN 12
RP SEQUENCE FROM N.A.
RA MEDLINE=20065165; PubMed=10597297;
RA Theodosiou A., Smith A., Gillieron C., Arkinstall S., Ashworth A.;
RT "MKP5, a new member of the MAP kinase phosphatase family, which
RT selectively dephosphorylates stress-activated kinases."
RL Oncogene 18:6981-6988(1999).
RN 13
RP SEQUENCE FROM N.A.
RA TISSUE=Brain, Lung, and Testis;
RC MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Heish F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Bork S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Richards S., Muzley K.C., Hale S., Garcia A.M., Gay L.J., Halysk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues Y., Sanchez A.,
RA Whiting M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

```

```

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Involved in the inactivation of MAP kinases. Has a
CC specificity for the MAPK1/MAPK2/MAPK3/MAPK4 subfamily.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB026436; BAB1668.1; -
DR EMBL: AF179212; AAD51857.1; -
DR EMBL: BC031405; AAH31405.1; -
DR HSSP: Q16828; MKP.
DR Genew: HGNC:3065; DUSP10.
DR GO: GO:0005737; Cytoplasm; TAS.
DR GO: GO:0005634; Cytoplasm; TAS.
DR GO: GO:0004721; Protein phosphatase activity; TAS.
DR GO: GO:0004721; Protein phosphatase activity; TAS.
DR GO: GO:0007254; P-JNK cascade; TAS.
DR GO: GO:0006950; Response to stress; TAS.
DR GO: GO:0006470; P-protein amino acid dephosphorylation; TAS.
DR InterPro: IPR000340; DS phosphatase.
DR InterPro: IPR008343; MAPK phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR00387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS: PRO1764; MAPKPHATASE.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANES 3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase: Nuclear protein.
FT DOMAIN 168 285 RHODANES.
FT DOMAIN 384 453 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 408 408 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 482 AA; 52642 MW; A8CB74ABF9498CD4 CRC64;

```

Query Match 13.7%; Score 469; DB 1; Length 482;
Best Local Similarity 35.9%; Pred. No. 2,2e-21;
Matches 107; Conservative 65; Mismatches 96; Indels 30; Gaps 7;

```

QY 27 LIDSRPFVEYNTSHILEAININCS-KLMKRLQODKVTITELIQ-HSARKYVDIDCSQK 84
DB 173 VIIDRPFVEYNTSHILEAININCS-KLMKRLQODKVTITELIQ-HSARKYVDIDCSQK 232
QY 65 VVVVYVDSQSO-VASLSSDCLTVLLGKLEKSPNSVHLLAGGFAEFSRCFPGLECGK----- 139
DB 233 IIVVDENTPESRVPQPHITVLESIKRGEKPEVLVKGGLSFKQNHNTLNDNSLQIOE 292
QY 140 -----STLVPCISQCLP-VANIGPRLIPNLYIGQGVVANKELIQNGIGY 187
DB 233 CREVGAGSASASSLLPPI-PTTPDLENALPILPFLFEDNEDADADLDTPMRLNIGY 350
QY 188 VVANS-----YTCKPDPFIESHFLRPVNDSCFKILFWLDSVDFIEKAKASNCVTV 242

```

```

DB 351 VINVTHTLPLHYEKGFLF-----NYKRLPATDSNKQMLRQYFEFAFERIEBAHQCGKGLLI 406
QY 243 HCLAGISRSATTAIAIYMKRMDNSLDEAYRFYKERTPTISPNNFPGQLLDYEKKIKN 300
DB 407 HCOAGVSRSAITVIAIYMKRMTMTDAYKFKVGRKPIISPNNFPGQLLEFEDLINN 464

RESULT 7
DUSA_MOUSE STANDARD; PRT; 483 AA.
AC Q9ESG0; Q9CZY9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase phosphatase 5) (MKP-5).
DE DUSP10 OR MKP5.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20515587; PubMed=11060451;
RA Matsuda K., Shima H., Kikuchi K., Matanabe Y., Matsuda Y.;
RT "Expression and comparative chromosomal mapping of MKP-5 gene DUSP10/Dusp10."
RL Cytogenet. Cell Genet. 90:71-74(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I., Saito T., Matsuda H., Gojobori T., Bono H., Kasukawa T., Saito R., Kadoya K., Okazaki H., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pezole G., Quackenbush J., Schraml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Maehiro T., Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barth G., Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.F., Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: Involved in the inactivation of MAP kinases. Has a specificity for the MAPK1/MAPK2/MAPK3/MAPK4 subfamily (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein + tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein + phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (by similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB037908; BAB17680.1; -.
DR EMBL: AK011995; BAB27966.1; -.
DR HSSP: Q16828; 1MKP.
DR MGD: MGI:1927070; Dusp10.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; Tyr_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; MAPKPHPTASE.
DR PRINTS: PR01764; MAPKPHPTASE.
DR SMART: SM00450; RHOD. 1.
DR PROSITE: PS50206; RHODANSE_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Nuclear protein.
FT DOMAIN 169 286 RHODANSE.
FT DOMAIN 386 456 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 409 409 PHOSPHOCYSTEINE INTERMEDIATE (BY SIMILARITY).
FT FT 48 48 A -> T (IN REF. 2).
FT FT 113 113 P -> S (IN REF. 2).
FT FT 386 386 R -> Q (IN REF. 2).
SQ SEQUENCE 483 AA; 52530 MW; 3551562355846f89 CRC64;

Query Match 13.7%; Score 468; DB 1; Length 483;
Best Local Similarity 35.6%; Pred. No. 2,5e-21;
Matches 106; Conservative 67; Mismatches 95; Indels 30; Gaps 7;

QY 27 LLDISRPVEVYNTSHLEAININCS-KLMKRLQDQKVLITLIQ-HSAKHYDIDCSQK 84
DB 174 VIIDCRPFMEYVNSKHIGAVHINCAIDISRRRLQCKITVLDISCRBGDSFKRTFSKE 233
QY 85 VVYVDDSSQDVASLSDCETVLVGLKESFNVSVHLLAGGFAPFSRCFPGLCGK----- 139
DB 234 IIVYDENTNTPSKSVYTPQPLVTLVLSLKKRGEPLVLKGLSFKQKHGKLNCSNLSLOE 293
QY 140 -----STLVPTCISOPCLP-VANIGPRLIPNLYLGGORDVNLKELLQONGIGY 187
DB 294 CREVGGAASASMLDQ--SVPTTPDIEANELTLPPLFLGNEDQAQDLDTMQRNLITGV 351
QY 188 VLNAS-----YTCPKDPFIPESHFLRPVNDSECEKILPWLKSVDFIEKAKSNGCVLV 242
DB 352 VINVTHTLPLHYEKGFLF-----NYKRLPATDSNKQMLRQYFEFAFERIEBAHQCGKGLLI 407
QY 243 HCLAGISRSATTAIAIYMKRMDNSLDEAYRFYKERTPTISPNNFPGQLLDYEKKIKN 300
DB 408 HCOAGVSRSAITVIAIYMKRMTMTDAYKFKVGRKPIISPNNFPGQLLEFEDLINN 465

RESULT 8
DUS6_MOUSE STANDARD; PRT; 381 AA.
AC Q9DBB1; Q9D7L4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase phosphatase 3) (MKP-3).
DE DUSP6 OR MKP3.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=liver, and Tongue;
RX MEDLINE=21085660; PubMed=11217851;

```

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi U., Fukuda S.,
RA Atakawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner T., Batcalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno A., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carrinci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hormann M., Hume D.A., Kamyta M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima M., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmung L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Raha S.A., Loggellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Bosak S.A., McMan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,
RA Villallon D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodchenko Y., Bouffard G.G.,
RA Whiting R.W., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley A.C., Grimwood J., Schmutz J., Myers R.W.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002)
CC -!- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
family (By similarity).
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
tyrosine + phosphate.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -!- Non-receptor class dual specificity subfamily.
CC -!- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
CC DR EMBL; AK005062; BAB23766.1; -
CC DR EMBL; AK009131; BAB26093.1; -
CC DR EMBL; BC003869; AAH03869.1; -
CC DR HSSP; Q16828; IMKP.
CC MGD; MGI:191485; Dusp6.
CC InterPro: IPR000340; DS phosphatase.
CC InterPro: IPR008343; MAPK phosph.
CC InterPro: IPR001763; Rhodanese-like.
CC InterPro: IPR000387; TYR phosphatase.
CC Pfam; PF00782; DUSPc; 1.
CC PRINTS; PR01764; MAPKPHPTASE.

DR SMART; SM00195; DUSPc; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE 3; 1.
DR PROSITE; PS00383; TYR PHOSPHATASE 1; FALSE_NEG.
DR PROSITE; PS50056; TYR PHOSPHATASE 2; 1.
DR PROSITE; PS50054; TYR PHOSPHATASE_DUAL; 1.
DR Hydrolase.
FT DOMAIN 30 148 RHODANSE.
FT DOMAIN 206 381 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 293 293 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
FT COMPACT 22 22 L -> G (IN REF. 1; BAB26093).
FT CONFLICT 34 34 W -> F (IN REF. 1; BAB26093).
SQ SEQUENCE 381 AA; 42407 MW; 7EALPBI54FAD2DA CRC64;
Query Match 13.6%; Score 466.5; DB 1; Length 381;
Best Local Similarity 33.7%; Pred. No. 2.2e-21;
Matches 112; Conservative 64; Mismatches 107; Indels 49; Gaps 8;
QY 19 LESTGEKVLIDSRPFVEVNTSHLEININCKMKRRRLQDQKVLITELIQ-SAKHKV 77
DB 27 LELGNERLLMDCKRPODELVSSESHIESAINALPIMLRLOKGNLPRALFTRCEDRDRF 86
QY 78 DIDC-SQKVVVYVYDQSSQD-VASLSQDPLTVLLGLKLEKFSNVYHLAGGFARFCFPG 135
DB 87 TRRCGTDTVVLVDENSDNMENNGSGSVGLLKLKDKGCAAFYLEGFSKFGQAFALH 146
QY 136 CEKSTIVTPVC-ISQCLPVANIG----- 158
DB 147 CE--TNIDGSCSSSPPLPLVGLGRLISSDSSDIESDLDRDPNSATSDSPLSNSQP 204
QY 159 --PTRLIPNLYAGCORVDYLNKEILQONGIGVYLNASTYCPK--DFPESHFLRPVNSF 215
DB 205 SFVELLPFLYLGCAQDNTNLDVLEFGKILNTVPLNPLPENNAGEFKYQIPLSDHW 264
QY 216 CEKILPWLKSDVFIEKASNGCYLVHCLAGISRSATIAIYIMKMDMSIDEAYRFX 275
DB 265 SQMLSQFPEALISFIDEARGKNCVYHCLAGISSVTVYVLMQKNTLSNNDAYDYLK 324
QY 276 EKRPTSPNPNFPGOLDYEXKIKKQGTASGP 307
DB 325 MKASNISPNFNGQLDPERTL---GLSSP 352
RESULT 9
DUS6_RAT
ID DUS6_RAT STANDARD; PRT; 381 AA.
AC 064346;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
DE phosphatase 3) (MKP-3).
GN DUSP6 OR MKP3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Neuron;
RX MEDLINE=96224012; PubMed=8626780;
RA Muda M., Boschart U., Dickinson R., Martinou J.C., Martinou I.,
RA Camps M., Schlegel W., Arkinstall S.;
RT "MKP-3, a novel cytosolic protein-tyrosine phosphatase that
RT exemplifies a new class of mitogen-activated protein kinase
RT phosphatase."
RL J. Biol. Chem. 271:4319-4326(1996).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96216487; PubMed=8631996;
RA Mourey R.J., Vega Q.C., Campbell J.S., Wenderoth M.P., Hauschka S.D.,

RA Krebs E.G., Dixon J.E.;
RT "A novel cytoplasmic dual specificity protein tyrosine phosphatase
RL J. Biol. Chem. 271:3795-3802(1996)."
CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
CC family. Implied in muscle and neuronal differentiation.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).
CC -----
CC EMBL: X94185; CAA63895.1; -;
CC EMBL: U42627; AAB06202.1; -;
CC HSSP: Q16828; 1MKP.
CC InterPro: IPR000340; DS_phosphatase.
CC InterPro: IPR008343; MAPK_phosph.
CC InterPro: IPR001763; Rhodanese-like.
CC InterPro: IPR000387; TYR_phosphatase.
CC Pfam: PF00782; DSPC; 1.
CC Pfam: PF00581; Rhodanese; 1.
CC PRINTS: PR01764; MAPKPHPTASE.
CC SMART: SM00195; DSPC; 1.
CC SMART: SM00450; RHOD; 1.
CC PROSITE: PS50206; RHODANASE_3; 1.
CC PROSITE: PS50383; TYR_PHOSPHATASE_1; FALSE_NEG.
CC PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
CC PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
CC KWL: HydroLase.
CC FT DOMAIN 30 148 RHODANASE.
CC FT DOMAIN 206 381 PROTEIN-TYROSINE PHOSPHATASE.
CC FT ACT_SITE 293 293 PHOSPHOCYSTEINE INTERMEDIATE (BY
CC SIMILARITY).
CC SQ SEQUENCE 381 AA; 42318 MW; C51LE0CB68F2868 CRC64;
Query Match 13.6%; Score 465.5; DB 1; Length 381;
Best Local Similarity 33.7%; Pred. No. 2.6e-21;
Matches 112; Conservative 64; Mismatches 107; Indels 49; Gaps 8;

RESULT 10
ID DUS4_HUMAN STANDARD; PRT; 394 AA.
AC Q13115; Q13524;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
DE phosphatase-2) (MKP-2) (Dual specificity protein phosphatase hwh2).
GN DUSP4 OR MKP2 OR VH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95221370; PubMed=7535768;
RA Guan K.-L., Butch E.;
RT "Isolation and characterization of a novel dual specific phosphatase,
RT HWH2, which selectively dephosphorylates the mitogen-activated
RT protein kinase.";
RL J. Biol. Chem. 270:7197-7203 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96198119; PubMed=8626452;
RA Chu Y., Solaki P.A., Khosravi-Far R., Der C.J., Kelly K.;
RT "The mitogen-activated protein kinase phosphatases PAC1, MKP-1, and
RT MKP-2 have unique substrate specificities and reduced activity in vivo
RT toward the ERK2 severmaker mutation.";
RL J. Biol. Chem. 271:6497-6501 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin, and Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loughellano N.A., Peters G.D., Abramson R.D., Mulhaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttrifield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases
CC ERK1 and ERK2.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: U21108; AAA85119.1; -
 DR EMBL: U48807; AAC50452.1; -
 DR EMBL: BC002671; AAH02671.1; -
 DR EMBL: BC014565; AAH14565.1; -
 DR HSSP: Q16828; IMKP.
 DR Genew; HGNC:3070; DUSP4.
 DR MIM; 602747; -
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0008330; F:protein tyrosine/threonine phosphatase acti. .; TAS.
 DR GO; GO:0000165; P:MAPKK cascade; TAS.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
 DR InterPro; IPRO00340; DS_phosphatase.
 DR InterPro; IPRO08343; MAPK_phosph.
 DR InterPro; IPRO01763; Rhodanese-like.
 DR InterPro; IPRO00387; TYR_phosphatase.
 DR Pfam; PR00782; DSPC; 1.
 DR Pfam; PR00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANES_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR HydroLase; Nuclear protein.
 FT DOMAIN 41 159 RHODANES.
 FT ACT_SITE 197 394 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 280 280 PHOSPHOCSTEINE INTERMEDIATE (BY SIMILARITY)
 FT CONFLICT 111 111 R -> G (IN REF. 2).
 FT SEQUENCE 394 AA; 42953 MW; 0603971759B6952E CRC64;

Query Match 13.0%; Score 445; DB 1; Length 394;
 Best Local Similarity 31.4%; Pred. No. 4.8e-20;
 Matches 120; Conservative 75; Mismatches 141; Indels 46; Gaps 14;

QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCKLMKRLQODKVLITELI--CHSAKH 76
 DB 39 LPSSG-GKCLLDRCRFLAHAGYILGVSIVNRCNTIVRR-AGSVSLQGLIPAEVRRAR 96
 QY 77 VIDDSOKVVYDSSQVAVSLSDCFILVLGLKEKF--NSVHLAGFAEFSRCFPG 134
 DB 97 IASGLYSAYIVYDERSPPAESLRDSTVSLVVAQLRRAERTDCLKGGYERFSSEYPE 156
 QY 135 LCEGKSTL-----VPTCISQP-----C-LPVANI-GPTRIILPNTYLGQRVNLKEL 179
 DB 157 FCSKTKALAIIPPVPSPATERLDLCSSCGTPIHDOGFPVEILLPIYLSSAHYAAARDM 216
 QY 180 IQQNGIGVLAASYCEKPDPTPESHFLRVVNDSPCEKILLPWLDKSVDFTEKAKANGC 239
 DB 217 LDALGITALLNVSSDCPN-HFEHGXYKCIPVENHNAIDISSWMEAEIYIDAVKDCRGR 275
 QY 240 VLVECLAGISRSATIALYIKRMDMSIDEAYRVKKEKPTISNPNFLQQLDYEEKIK 299
 DB 276 VLVHQAQGISRSATICLAYLMKKRVRIEAFVKKRSTISNFSFMQQLQFESQVYL 335
 QY 300 NOTGASGPKSKLKLHLHKEPNEPVAVSEGGOKSETEPLSPCADATSEAAAGRPVYPAS 359
 DB 336 ATSCAAEAAAS-----PSGPL-----REKGTATP-----TSQVFSEFPV-SVG 373
 QY 360 VPSVPSQPSLLEDSPLVQALS 381
 DB 374 VHSAPSSLPYL--HSPITTSFS 393

RESULT 11
 ID DUS4 CHICK STANDARD; PRT; 375 AA.
 AC Q9PW71;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
 DE phosphatase-2) (MKP-2).
 GN DUSP4 OR MKP2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White leghorn;
 RX MEDLINE=20379359; PubMed=10918612;
 RA Fu S.-L., Wana A., Vogt P.K.,
 RT "Identification and characterization of genes upregulated in cells
 RT transformed by v-Jun".
 RL Oncogene 19:3537-3545(2000).
 CC -1- FUNCTION: Regulates mitogenic signal transduction by
 CC dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
 CC and ERK2 (By similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF167296; AAD46656.1; -
 DR HSSP: Q16828; IMKP.
 DR InterPro; IPRO00340; DS_phosphatase.
 DR InterPro; IPRO08343; MAPK_phosph.
 DR InterPro; IPRO01763; Rhodanese-like.
 DR InterPro; IPRO00387; TYR_phosphatase.
 DR Pfam; PR00782; DSPC; 1.
 DR Pfam; PR00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANES_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR HydroLase; Nuclear protein.
 FT DOMAIN 25 143 RHODANES.
 FT ACT_SITE 178 375 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 261 261 PHOSPHOCSTEINE INTERMEDIATE (BY SIMILARITY)
 FT SEQUENCE 375 AA; 41052 MW; 179290DC2BEEEF1 CRC64;

Query Match 13.0%; Score 444; DB 1; Length 375;
 Best Local Similarity 31.1%; Pred. No. 5.1e-20;
 Matches 123; Conservative 73; Mismatches 152; Indels 48; Gaps 13;

QY 4 EMIGQIVTERVALLBSGTEKVLIDSRPVEYNTSHILEAININCKLMKRLQODKV 63
 DB 9 EMEGSAL--RLVGRERESGRCCLLDRCRFLAHAGYILGVSIVNRCNTIVRR-AGAV 65
 QY 64 LITELI--CHSAKHKVIDDSOKVVYDSSQVAVSLSDCFILVLGLKEKF--NSVH 119
 DB 66 SLQGLPRAEGEVRRARLAAGYLTAIVVLYDERSPPAESLRDSTVSLVVAQLRRAERDADR 125

```

QY 120 LLAGFAEBSRCFGLCEGKSTLVPTICISOP-----C-LPVANI-GPTRILPN 165
DB 126 LLAGVIRFSEKPEFAKTKTL--SISPSASAESIDLGSSCGTPPLHDQGPVEILPE 183
QY 166 LYLGCGRDVLNKLIDQNGIGVYLNASYTCPEKPDPIESHFLRVPVNDSPCEKILPWLDK 225
DB 184 LYLGSATYHARRMDLALGITALLNVSSDCPN-HFEHYKYKCIPIEDNHKALISSWFE 242
QY 226 SVDFIEKAKASNCVILVHCLAGISRSATIAIYIMKMDNSLDEAYRFVKEKPTISPNF 285
DB 243 AIREYDSVEKCCGRVLVHCOAGISRSATICLAYIMMKRVKLEKAEFEVQORSIISPNR 302
QY 286 NPLGOLLIDYKTKIKNOGTAGSPKSKLKLHLEKNEVPVAVSEGGQSETPPLSPCADSA 345
DB 303 SFMGOLLQFESQVLTASCAVEAAS-----PSGTLRERGRATSTP----- 341
QY 346 TSEAGQRPVHPASVPSVPSQPSLSDPSLVQALS 381
DB 342 TSQFVFSFPV-SVGVAHPSSLPYL--HSPITTS 374

RESULT 12
DUS4_RAT STANDARD; PRT; 395 AA.
AC 062757;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
DE phosphatase-2) (MKP-2).
GN DUSP4 OR MKP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RF SEQUENCE FROM N.A.
RC TISSUE=Phenochromocytoma;
RX MEDLINE=95301550; PubMed=7782322;
RA Maira-Press A., Rim C.S., Yao H., Roberson M.S., Stork P.J.S.;
RT "A novel mitogen-activated protein kinase phosphatase. Structure,
RT expression, and regulation."
RT J. Biol. Chem. 270:14587-14596(1995).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
CC and ERK2 (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed at moderate levels in nearly all
CC tissues and cells including brain, spleen, and testes with the
CC higher expression in the heart and lung and lower expression in
CC skeletal muscle and kidney. Undetectable in liver. Expressed in
CC many areas of the brain with very strong expression in the
CC hippocampus, piriform cortex, and the suprachiasmatic nucleus.
CC -1- INDUCTION: By mitogens and by stress.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U23438; AAC52493.1; -.

```

```

DR HSP; Q16828; MKP.
DR InterPro; IPR000340; D5_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00782; D5PC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; D5PC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase; Nucleic protein.
FT DOMAIN 42 160 RHODANES.
FT DOMAIN 198 395 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 281 281 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT ACT_SITE 281 281 SIMILARITY).
SQ SEQUENCE 395 AA; 43187 MW; A90EFPD378A050FD CRC64;

Query Match 12.9%; Score 439.5; DB 1; Length 395;
Best Local Similarity 30.5%; Pred. No. 1e-19;
Matches 118; Conservative 71; Mismatches 147; Indels 51; Gaps 12;

QY 16 VALLESGETKLVLLIDRPFVEYNTSHLEAININGSKMKRLQDQKVLITELI--QHSA 73
DB 38 LGLLSGG--KCLLLDPRPLASAGYIRGSVNVRCITIVRR-AKGSVSLQILLPAEEV 94
QY 74 KHKVDIDCSQKVVVYVYDQSSQDVASLSSDCLFVTLGLKLEKSF--NSVHLLAGFAEBSRC 131
DB 95 RARLRGLYSANIVYVDERSPRAESLREDSTSLVQALRRNERTDICLKGKEFFSE 154
QY 132 FPGLCGKSTLVPTICISOPCLPVANI-----GPTRIIPNLYLGCQRDV 174
DB 155 YEFCEKTKAL--AAPRPVPSTNESLIDGSSCGTPPLHDQGPVEILPFLVLSAYHA 212
QY 175 LNKELIQNGIGVYLNASTYTCPEKPDPIESHFLRVPVNDSPCEKILPWLDKSDVLEKX 234
DB 213 ARDDMDLALGITALLNVSSDCPN-HFEHYKYKCIPIEDNHKADISSWMEALEYIDAVK 271
QY 235 ASNGCVLVHCLAGISRSATIAIYIMKMDNSLDEAYRFVKEKPTISPNFPLQLDLY 294
DB 272 DCRGRVLVHCOAGISRSATICLAYIMMKRVLEKAEFEVQORSIISNFSMGQLLP 331
QY 295 EKKIKNOGTAGSPKSKLKLHLEKNEVPVAVSEGGQSETPPLSPCADSATSEAGQRP 354
DB 332 ESQVLTSCAEEAAS-----PFGP--LRERKATPTP-----TSQFVFSFP 370
QY 355 VHPASVPSVPSQPSLSDPSLVQALS 381
DB 371 V-SVGVAHPSSLPYL--HSPITTS 394

RESULT 13
DUS1_MOUSE STANDARD; PRT; 367 AA.
AC P28563;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE 3CH134) (Protein-tyrosine phosphatase ERP).
GN DUSP1 OR PTPN10 OR MKP1 OR 3CH134 OR PTPN16.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RF SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=92158357; PubMed=1741163;

```

RA Charles C.H., Adler A.S., Lau L.F.;
 RT "cDNA sequence of a growth factor-inducible immediate early gene and
 RT characterization of its encoded protein.";
 RL Oncogene 7:187-190(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93360956; PubMed=8355678;
 RA Noguchi T., Metz R., Chen L., Mattei M.-G., Carrasco D., Bravo R.;
 RT "Structure, mapping, and expression of erp, a growth factor-inducible
 RT gene encoding a nontransmembrane protein tyrosine phosphatase, and
 RT effect of ERP on cell growth.";
 RL Mol. Cell. Biol. 13:5195-5205(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Czech II; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stedman M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughell N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyil S.W.,
 RA Villalón D.K., Mozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska M., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=94037096; PubMed=8221888;
 RA Sun H., Charles C.H., Lau L.F., Tonks N.K.;
 RT "MKP-1 (3CH14), an immediate early gene product, is a dual
 RT specificity phosphatase that dephosphorylates MAP kinase in vivo.";
 RL Cell 75:487-493(1993).
 CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
 CC kinase ERK2 on both Thr-183 and Tyr-185.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- INDUCTION: By growth factors.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X61940; CAA43944.1; -;
 CC EMBL; S64851; AAB27882.1; -;
 CC EMBL; BC006967; AAH06967.1; -;
 CC PIR; A54681; S24411.
 CC HSSP; Q16828; IMKP.
 CC MGD; MGI:105120; Dusp1.
 CC InterPro; IPR000340; DS phosphatase.
 CC InterPro; IPR008343; MAPK_phosph.
 CC InterPro; IPR001763; Rhodanese-like.
 CC InterPro; IPR000387; TYR_phosphatase.

DR Pfam; PF00782; DSpC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; NAKRPHPTASR.
 DR SMART; SM00195; DSpC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PSS0206; RHODANSE_3; 1.
 DR PROSITE; PSS0083; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase; Cell cycle.
 FT DOMAIN 20 137 RHODANSE
 FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE.
 FT MUTAGEN 258 258 C->S: LOSS OF ACTIVITY.
 SQ SEQUENCE 367 AA; 39369 MW; 50B5F90EBBD19AB CRC64;
 Query Match 12.7%; Score 433; DB 1; Length 367;
 Best Local Similarity 31.6%; Pred. No. 2.3e-19;
 Matches 115; Conservative 54; Mismatches 151; Indels 44; Gaps 8;
 QY 15 LVALLSGTEKVLIDSRPPEVNTSHILEAININGSKLMKRLQODKVLITELIOHSAK 74
 DB 13 LRALLRGAAGCILLDCRSFFAFNAGHIGSVVRSTIVRRRAKMGLEHIVPAEIR 72
 QY 75 HKYDIDCSQRYVYVYDSSQDVASLSDPELTVLGL--EKSRNSVHLLAGFAEBSRCF 132
 DB 73 GRLLAGYAHVAVLLDEKRSASLDGAKRDGTLAAGALCREARSTOVFFLGYEAFSASC 132
 QY 133 PGICGEGSTLVPCISGP-----CLVANI-----GPTRLPLNYGCGQDVL 175
 DB 133 PELCSKST--PGLSLPLSTSVSDASGSCSCSTPLXYOGFVELLSFLYIGASVHAS 190
 QY 176 NKEILQNGIGYLVNAYTCPEKDPFIPESHFLRPVNVDSFCEKILPWLDRKSVDFIERAKA 235
 DB 191 RKMULDALGITLALINVSANCPN-HFEGHYQKSIPEVDNKHADISSWFNAIDFIDSKD 249
 QY 236 SNGCVLVHICLAGISRSKATIALIATIMKMDSLDEARFVYEKPTISPNPNFGQLDYE 295
 DB 250 AGGRVFVHCQAGISRSATLICLAIVLMRTNRVYKLDFAEFVQRRSITISPNFSFGQLQFE 309
 QY 296 KIKIKNGASGSPSKULHLLEKNEPVPVASEGQGETPLSPCCADSATSEAAQGPV 355
 DB 310 SQV-----LAPHCS-----AEKSPNAAVLDRGISTTVTFNPNVS-----IPV 347
 QY 356 HPAS 359
 DB 348 HPTN 351
 RESULT 14
 DUS1_RAT
 ID DUS1_RAT STANDARD; PRT; 367 AA.
 AC 064623;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
 DE Cl100).
 GN DUSP1 OR Cl100.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognath1; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Lung;
 RA Mada M., Schlegel W., Arkininstall S.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
 CC kinase ERK2 on both Thr-183 and Tyr-185.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.

CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
phosphate.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X64004; CAA58828.1; -.
DR PIR: S52265; S52265.
DR HSSP: Q16828; 1MKP.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSEP; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR SMART: SM00195; DSEP; 1.
DR SMART: SM00450; RHOD; 1.
DR SMART: SM00206; RHODANES; 3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase; Cell cycle.
FT DOMAIN 20 137 RHODANES.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE (BY
SIMILARITY).
SQ SEQUENCE 367 AA; 39541 MW; 5112ADF290499139 CRC64;
Query Match 12.4%; Score 423; DB 1; Length 367;
Best Local Similarity 31.3%; Pred. No. 9.4e-19;
Matches 115; Conservative 54; Mismatches 148; Indels 50; Gaps 9;
QY 15 LVALLSGTEKVLIDSRPFEVNTSHLEAININCSKMKRRLLQ-----QDKVLTLELI 69
DB 13 LRLALLERAQCLLDRCRFFAFNAGIVSVVVRSTIYRRKAKMGLEHIVPTLEL- 71
QY 70 QHSARKHVDIDCSQKVVVYVDOSSQDVASISSDCEFLVLLGKL--EKSPNSVHLLAGGFAR 127
DB 72 ----RGRLLAGAYHAVVILDERSAALDGAKRGGTLAAGALCREARSTGVFLGGGYEA 127
QY 128 FSRCPFGLCGSK-----STVPTCISQPCLPVANI-----GPTRLIPMLYIGCOR 172
DB 128 FSASCEBELCSKQSTPMGLSLPLSTVSDSAESQSSCSPTLYDQGGVELSLFLYGSAY 187
QY 173 DVLNKLIIQONGIGVYVLAAGYTCKPDPFIPESHFLRPVNVDSCEKILPLDLSDVPEIEK 232
DB 188 HASRKMDLALGITALINVSANCPN-HFBSHYQIKSIPEDNKADISSENFENAIPTDS 246
QY 233 AKASNCVIVHCLAGISRSATIIAYIMKMDSLDEAYRFVKEKPTISPNFNFLGOLL 292
DB 247 IKDAGRVFHCQAGISRSATICLAYIMRTNRKYLDEAFVQGRSIIISPNFNGOLL 306
QY 293 DYKAKTKNQTGASGPKSKLKLHLEKNEPVPVBSGGQSETPPLSPCCADSAITSAAGQ 352
DB 307 QFSSQV-----LAFHCS-----AEGSPAMAVLDKRGSTTVTFNFPVS----- 344
QY 353 RPYHPAS 359
DB 345 IPVHPTN 351

AC P28562;
DT 01-FEB-1992 (Rel. 24, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE CL100) (Dual specificity protein phosphatase hVH1).
GN DUSP1 OR PRPNI1 OR MKP1 OR CL100 OR VH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Forebrain;
RX MEDLINE=93024952; PubMed=1406996;
RA Keyes S.M., Emalie E.A.;
RT "Oxidative stress and heat shock induce a human gene encoding a
RT protein-tyrosine phosphatase.";
RL Nature 359:644-647 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshynski S., Carninci P., Pange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mollay S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rodriguez Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scheraga A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
CC kinase ERK2 on both Thr-183 and Tyr-185.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- INDUCTION: By oxidative stress and heat shock.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68277; CAA48338.1; -.
DR EMBL: BC022463; AAH22463.1; -.
DR PIR: S29090; S29090.
DR HSSP: Q16828; 1MKP.
DR Genew; HGNC:3064; DUSP1.
DR MIM; 600714; -.
DR GO; GO:0004726; P:non-membrane spanning protein tyrosine phos. . .; TAS.
DR GO; GO:0006979; P:response to oxidative stress; TAS.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.

DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES 3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE 1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; Cell cycle.
FT DOMAIN 20 137 RHODANES.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE (BY
SIMILARITY).
SQ SEQUENCE 367 AA; 39297 MW; 11BD1D39A9FCD51F CRC64;

Query Match 12.3%; Score 420; DB 1; Length 367;
Best Local Similarity 31.2%; Pred. No. 1.4e-18;
Matches 113; Conservative 56; Mismatches 159; Indels 34; Gaps 7;

QY 15 LVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRLQODKVLITELIQHSK 74
DB 13 LRALLGERAAQCLLDRCSPFAFNAGHAGSVNVSFTIVRRRAKAMGLEHIVNMLR 72
QY 75 HKVDIDCSQKVVVDQSDVASLSDCFLTVLQKL--EKSPNSVHLLAGPAEFRCF 132
DB 73 GRLLAGAYHAYVLLDERSAALDGAARDGTLLALAGALCREARAQVFFLKGGYENFSASC 132
QY 133 PGLCEGK-----STLVPTCISQPLPVANI-----GPTRLIPNLVLGQORDVLNK 177
DB 133 PELCSKQSTPMGLSLPLSTSVPSDAESGSCSTPLYDQGVLEILPFLYLGSAYHASRK 192
QY 178 ELIQNGIGYVLNMSYTPKPEDFIPESHFLKVPVNDSCFKILFWLDSVDFIEKAKASN 237
DB 193 DMLDALGITALINVSANCPN-HFEGHYQYKSIPEVDNHKADISSWFNEAIDFIDSIKNAG 251
QY 238 GCVLVHCLAGISRSATIAIATYMKRMDMSLDEAYRFVKEKRPRTISPNNFPLGOLLIDYEKK 297
DB 252 GRVFVHCQAGISRSATICTATYMKRTNRVYKLDPAEFVQKRSIISPNSFMQQLQFESQ 311
QY 298 IKNQTGASGPKSKLTLHLKRPNEPVAVSQGQKSET---PLSPPCADSATSEAAQQR 353
DB 312 V-----LAPHCS-----AEAGSPAMAVLDRGTSTTVFNFPVSIIVHSTNSALSTYQS 359
QY 354 PV 355
DB 360 PI 361

Search completed: June 21, 2004, 13:21:07
Job time : 16.4416 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 20:53:18 ; Search time 4983.03 Seconds
(without alignments)
3985.196 Million cell updates/sec

Title: US-10-029-345A-109
Perfect score: 3418
Sequence: 1 MAHEMGTQVTERVALLE.....LGRVSGSSFSGSMETIEVS 665

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q/cgnt 1/USPTO.spool/US10029345/rnat 21062004.164146.29039/app.query.fasta.1.1294
-DB=EST -CPMT=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029345.@CGN 1.1 3931.@rnat 21062004.164146.29039 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_hiv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1553	45.4	1060	12	BM546940
2	1412	41.3	881	13	BQ30140 AGENCOURT
3	1397.5	40.9	883	13	BQ45892 AGENCOURT
4	1375.5	40.2	974	13	BQ51695 AGENCOURT
5	1375	40.2	859	13	BQ33499 AGENCOURT
6	1369.5	40.1	898	13	BQ72165 AGENCOURT
7	1360	39.8	893	12	BG482429 AGENCOURT
8	1327	38.8	2207	11	BC038231 Homo sap1
9	1306	38.2	862	13	BQ770036
10	1216	35.6	769	13	BU704078
11	1180	34.5	836	12	BI821804
12	1148	33.6	715	14	CF727177
13	1102.5	32.3	920	10	BE897795
14	1086.5	31.8	682	12	BI16954
15	1081	31.6	656	14	CF532917
16	1076	31.5	639	14	CF742387
17	1075	31.5	891	13	BQ670989
18	1073.5	31.4	946	12	BI657528
19	1071.5	31.3	660	12	BI917706
20	1065	31.2	983	12	BG15188 602344051
21	1059.5	31.0	790	12	BI661614
22	1047	30.6	663	13	BK479029
23	1029.5	30.1	752	13	BQ432082
24	1006	29.4	650	10	BF472046
25	979	28.6	621	9	AV703072
26	978.5	28.6	924	14	CA474739
27	972.5	28.5	743	14	CF727802
28	970	28.4	693	9	AV701628
29	964.5	28.2	739	28	AZ850283
30	964.5	27.4	602	10	AM847426
31	931	27.2	595	10	AM847426
32	925	27.1	631	28	BH039241
33	911	26.7	1067	10	BF135687
34	891	26.1	3325	11	AK035652
35	872	25.5	792	13	BU946569
36	839.5	24.6	586	14	CF104291
37	828	24.2	723	13	BU708175
38	823.5	24.1	516	29	AY413655
39	804.5	23.5	516	29	AY413657
40	772.5	22.6	683	12	BM942528
41	770	22.5	619	29	CG583869
42	769	22.5	579	10	BF582366
43	768.5	22.5	594	14	CF727694
44	762	22.3	512	10	BF815601
45	749.5	21.9	538	10	AW909957

ALIGNMENTS

RESULT 1
LOCUS BM546940
DEFINITION AGENCOURT_6491295 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5723711
ACCESSION BM546940
VERSION 5', mRNA sequence.
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1060)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strassberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
[found through the I.M.A.G.E. Consortium/LLNL at:](http://image.llnl.gov)
<http://image.llnl.gov>
 Plate: LAM12711 row: m column: 24
 High quality sequence stop: 737.
 Location/Qualifiers

1. 1060
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5723711"
 /lab_host="DH10B"
 /clone_11b="NIH MGC 125"
 /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
 Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
 of three ovaries, from females ranging in age from 38 to
 49 yo. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 2.1 kb, insert size range 1-3.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 036."

ORIGIN

Alignment Scores:

Pred. No.: 2,95e-125 Length: 1060
 Score: 1553.00 Matches: 317
 Percent Similarity: 93.91% Conservative: 7
 Best Local Similarity: 91.88% Mismatches: 15
 Query Match: 45.44% Indels: 7
 DB: 12 Gaps: 2

US-10-029-345a-109 (1-665) x BM546940 (1-1060)

QY 163 LeuProAsnLeuTyrlenglyCysGlnarGpValLeuAsnlyGluLeuIleGln 182
 Db 6 ATTCCCAATCTTATCTTGCTGCAGCAGATGCTCCACAAAGAGCTGATCAGAG 65
 QY 183 AsnGlyIleGlyTyrlValLeuAsnAlaserTyrlThyrProlyPProAspPheilePro 202
 Db 66 AATGGGATTGTTAGTGTAAATCCCAACCAATCTGTCACAAAGCCTGACTTATCCCC 125
 QY 203 GluSerHisPheLeuAlaValProValAsnAspSerPheCysGluTyrlLeuAspPro 222
 Db 126 GAGCTCTCATTTCTCTGCTGCTGCTGATGACAGCTTTTGAGAAAATTTTCCCGTGG 185
 QY 223 LeuAspLySerValAspPheIleGluAlaValAsnAlaserGlyCysValLeuVal 242
 Db 186 TTGGCAATATGATGATTTTATGAGAAAGCAAAAGCCTCCATGATGATGTTCTAGTG 245
 QY 243 HisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlMetLysArg 262
 Db 246 CACTGTTTATGCTGGAGATCTCCCGCTCCGACCATGCTATGCTCATCATGAGAGG 305
 QY 263 MetAspSerLeuAspGluAlaTyrlArgPheValIleGluAlaArgProThrIleSer 282
 Db 306 ATGACAGATGCTTTAGATGAGAGCTTCAAGATTGTAAGAAAAGAAAGCTACTATATCT 365
 QY 283 ProAsnPheAsnPheLeuGlyIleLeuLeuAspTyrlGluTyrlIleLysAsnGlnThr 302
 Db 366 CCAAACTTCAATTTCTGGGCCAACTCTGACATGAGAGAGATTTAAGACCAAGACT 425
 QY 303 GlyAlaSerGlyProlySerLySerLyLeuLeuHISleuGluTyrlProAsnGluPro 322

Db 426 GZAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGCACCTGGAGAAACCAATGAACCT 485
 QY 323 ValProAlaValSerGluGlyGlyIleLysSerGluThrProLysSerProProCysAla 342
 Db 486 GTCCCTGCTGCTCAGAGAGGTGACAGAAAAGCAGAGAGCCCTCATCTCAGCTGTGCC 545
 QY 343 AspSerAlaThrSerGluAlaAlaGlyIleLysArgProValHisProAlaSerValProSer 362
 Db 546 GACTCTGACTCCTCAGAGAGCAGAGCAAAAGGCCCTGTGATCCCGC-----AGC 596
 QY 363 ValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeuSerGly 382
 Db 597 GTACCACTGCTGAGCCGTGCTGTTAGAGAGACAGCCCGCTGTGTACAGCGCTCATGTGG 656
 QY 383 LeuHisLeuSerAlaAspArgLeuGluAspSerAsnLyLeuLysArgSerPheSerLeu 402
 Db 657 CTGCACCTGCTCCGACAGACGCTGGAAGACAGCAATAGCTCAAGCGTTCTTCTCTCG 716
 QY 403 AspIleLysSerValSerTyrlSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSer 422
 Db 717 GATATCAAAATCAGTTTCAATATTCAGCCAGCATGAGCATCTTTCATAGGCTTCTCTCA 776
 QY 423 SerGluAspAlaLeuGluTyrlTyrlLysProSerThrThrLeuAspGlyThrAsnLysLeu 442
 Db 777 TCAGAGATGCTTTGGAATTAATAAACCCTTCACTCTGAGATGGACCAACCAACTA 836
 QY 443 CysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerProAspLys 462
 Db 837 TGGCAATTTCTCCCGTANTCANAACTATCGACAGACAGACTCCGAAACAGATCT- GATNAG 895
 QY 463 GluGluAlaSerIleProLysLysLeuGlnThrAlaArgPro-SerAspSerGlnSerly 482
 Db 896 GAGAGAGCCAGCATCCCGACAGAGTGCAGACCCCGAGGCTTTCACAGACCAAGAGCA 955
 QY 482 SarG-LeuHisSerValArgThrSerSer---GlyThrAlaGlnArgSerLeuLys 501
 Db 956 GCGGATTCATTCAGTANTGAGAACCCAGCAGCAGAGTGCGACCGCCAGAGANGTCCCTTT 1015
 QY 501 erProLeuHis 504
 Db 1016 TTATCTCTCAC 1026

RESULT 2
 BQ30140
 LOCUS
 DEFINITION
 AGENCOURT_8923732 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6460964
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 881)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE
 JOURNAL
 COMMENT
 Unpublished (1999)
 Contact: Robert Strassberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
[found through the I.M.A.G.E. Consortium/LLNL at:](http://image.llnl.gov)
<http://image.llnl.gov>
 Plate: L1CM2648 row: h column: 21
 High quality sequence stop: 676.
 Location/Qualifiers

1. 881
 /organism="Homo sapiens"
 /mol_type="mRNA"

/db_xref="taxon:9606"
 /clone="IMAGE:6460964"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH MGC 101"
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 4,98e-113 Length: 881
 Score: 1412.00 Matches: 285
 Percent Similarity: 98.62% Conservative: 0
 Best Local Similarity: 98.62% Mismatches: 1
 Query Match: 41.31% Indels: 3
 DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x BQ930140 (1-881)

QY 266 SerLeuAapGluAlaTyrArgPheValIyGluLysArgProThrIleSerProAapPhe 285
 DB 3 TCTTTAGATGAGGCTTACAGATTGTGTGAAAGAAAAAGACCTTATATCTCCAAACTTC 62
 QY 286 AapPheLyuGlyGlnLeuAapTyrGluLysArgIleLysAangInThrgIyAlaSer 305
 DB 63 AATTTCCTGAGCCCAACTCTGAGCTATGAGAGAGATTAAGAACCAAGACTGGAGCAATCA 122
 QY 306 GlyProLysSerIyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 325
 DB 123 GGGGCAAGAGCAAACTCAAGCTGCTCACTGGAGAGCAATATGAATGAACTGCTCCGCT 182
 QY 326 ValSerGluGlyGlnLysSerGluThrProLysSerProProCysAlaAapSerAla 345
 DB 183 GTCTCAGAGGGGTGACGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 242
 QY 346 ThrSerGluAlaIyGlnArgProValHisProAlaSerValProSerValProSer 365
 DB 243 ACCCTAGAGGAG 302
 QY 366 ValGlnProSerLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLys 385
 DB 303 GTGCAAGCCGTGCTGTTAG 362
 QY 386 SerAlaAapArgLeuGluAapSerAapLysLysLysLysLysLysLysLysLysLys 405
 DB 363 TCCGAGAGACAGCTGAG 422
 QY 406 SerValSerIySerAlaSerMetAlaLysSerLeuHisLysLysLysLysLysLys 425
 DB 423 TCAGTTTCATTCAGCAG 482
 QY 426 AlaLeuGluTyrIyLysProSerThrThrLeuAapGlyThrAanLysLysLysLys 445
 DB 483 GTTTTGAATATCAACAACCTTCACATCTGATGGAGCAACAAGCATATGCAAGTTTC 542
 QY 446 SerProValGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 465
 DB 543 TCCCTCTTCAAGAACTATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 602
 QY 466 SerIleProLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 485
 DB 603 AGCATCCCAAG 662
 QY 486 SerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLysLysLysLysLys 505
 DB 663 TCGGTCAAG 722

QY 506 SerGlySerValGluAapAapTyrHisSerPheLeu-PheGlyLeuSerThrSergI 525
 DB 723 AGTGGAGAGGTGGAG 782
 QY 525 ngInHisLeuThrLysSerAlaGlyLeu-GlyLeuLysGlyTyrPHisSerAapIleLeu 545
 DB 783 GCAGCACTTCAGCAATCTGCTGCTGGGGCTTAAAGGCTGAGCATCGATATCTTGG 842
 QY 545 la-ProGlnThrSerThrPro 551
 DB 843 CCCCCCAGACCTTACCCCT 863

RESULT 3

BQ945892 883 bp mRNA linear EST 21-AUG-2002
 LOCUS AGENCOURT_8926314 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6463045
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ945892
 VERSION BQ945892.1 GI:22361370
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 883)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLCM2653 row: 0 column: 14
 High quality sequence stop: 672.
 Location/Qualifiers
 1. 883
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6463045"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

FEATURES

source

1. 883
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6463045"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 9.34e-112 Length: 883
 Score: 1397.50 Matches: 286
 Percent Similarity: 97.29% Conservative: 1
 Best Local Similarity: 96.95% Mismatches: 8
 Query Match: 40.89% Indels: 3
 DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x BQ945892 (1-883)

QY 266 SerLeuAapGluAlaTyrArgPheValIyGluLysArgProThrIleSerProAapPhe 285
 DB 3 TCTTTAGATGAGGCTTACAGATTGTGTGAAAGAAAAAGACT-ACATATCTCCAA--CTT 59
 QY 286 AapPheLyuGlyGlnLeuAapTyrGluLysArgIleLysAangInThrgIyAlaSer 305

```

Db      60 CATTTCCTGGGCAACTCTGAGATATGAGAAAGATTAAAGAACAGATGAGCATCA 119
Qy      306 GTPYPLYSerIySleuIySleuHISleuGluIySProAnGluProValProAla 325
Db      120 GGGCCAAAGACAACTCAAGCTGCTCACCTGAGAGCAATTAACCTGCTCT 179
Qy      326 ValSerGluGlyGlyGlySleuSergIyThrProIleuSerProProCyAlaAspSerAla 345
Db      180 GTCTCAGAGGGGTGACAGAAAGGAGAGCCCTCTCATGTCACCTGTCACACTGCT 239
Qy      346 ThSerGluAlaAlaGlyGlnArgProValHisProAlaSerValProSerValProSer 365
Db      240 ACCCTCAGAGGACGAGAGCAAAAGCCCGTGCATCCGCGACGCTGCCAGCGTCCAGC 299
Qy      366 ValGlnProSerIleuGluIySleuSergProIleuValGlnAlaSerGlyIleuHISleu 385
Db      300 GTGCAGCGCTGCTGTTAGAGGAGCAGCCCGCTGCTGTAAGGCGCTCAAGGGGCTGCACCTG 359
Qy      386 SerAlaAspArgIleuGluIySleuSergIySleuIySArgSerPheSerIleuAspIleu 405
Db      360 TCCGCAAGACGCTGAGAGCAAGCAATTAAGCTCAAGCGTTCCTCTCTGATATCAAA 419
Qy      406 SerValSerIySergIySergIySergIySergIySergIySergIySergIySergIySerg 425
Db      420 TCAGTTTCATATTTCAGCCGAGATGCGACATCTTACATGAGCTTCTCTCATCAGAAAGAT 479
Qy      426 AlaIleuGluIyTyTyIySProSerThrIleuAspGlyThrAsnIySleuCySglnPhe 445
Db      480 GTTTGGAAATCTAACAACTTCACATCTGATGGAGCAACAAAGATGCCAGTTTC 539
Qy      446 SerProValGlnIleuSergIyGlnIyThrProGluIySergIySergIySergIySergIy 465
Db      540 TCCCTGTTTCAGAACTATCGAGACAGATCCGGAACAGCTCTATTAAGAGAGAGGC 599
Qy      466 SerIleuProIySleuGlnIyThrAlaArgProSerAspSergIySergIySergIySergIy 485
Db      600 AGCATCCCAAGAGAGCTGACAGCCGCGCTTCAGACAGCAGCAGAGAGAGGATTCAT 659
Qy      486 SerValArgThrSerSerSergIyThrAlaArgSerIleuSergIySergIySergIySergIy 505
Db      660 TCGGTCAAGAAACAGACAGATGCGACCGCCAGAGATCCCTTTATTCCTCATCTCATCGA 719
Qy      506 SerGlySerValGlnIySleuSergIySergIySergIySergIySergIySergIySergIy 525
Db      720 AGTGGAGAGCTGAGAGCAATTAACAGACAGACCTTCCTTCGCGCTTCACACAGCAG 779
Qy      526 GlnHISleuThrIySergIySergIySergIySergIySergIySergIySergIySergIy 545
Db      780 CAGCACTCAGAGAGCTGCTGCGCTGGGGCTTAA-GGCTGGCAGCTGGATATCTGGGCC 838
Qy      546 ProGlnThrSerThrProSerIleuThrSerSergIyTyTyPheAla 560
Db      839 CCCCAAGCTTACCCCTTCCTGACAGCAGCTGGATATTTGCC 883

RESULT 4
LOCUS   B0951695          974 bp      mRNA      linear      EST 21-AUG-2002
DEFINITION  AGENCOURT 8947130 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6461668
ACCESSION  B0951695
VERSION    B0951695.1 GI:22367173
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (bases 1 to 974)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
COMMENT

```

FEATURES

source

Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNC2650 row: F column: 05
 High quality sequence stop: 617.
 Location/Qualifiers

```

1..974
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6461668"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_101"
/note="Organ: Lung; Vector: pOTB7, Site 1: EcoRI, Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

ORIGIN

Alignment Scores:

```

Pred. No.: 9,37e-110      Length: 974
Score: 1375.50      Matches: 291
Percent Similarity: 91.82%      Conservative: 1
Best Local Similarity: 91.51%      Mismatches: 18
Query Match: 40.24%      Indels: 8
DB: 13      Gaps: 3

```

US-10-029-345A-109 (1-665) x B0951695 (1-974)

```

Qy      266 SerIleuAspGluAlaIyTyArgPheValIySglnIySArgProThrIleSergProAnPhe 285
Db      3 TCTTTAGATGAAGCTTTCAGATTGTGTAAGAAAGAAAGACCTATATCTCCAACTTC 62
Qy      286 AsnPheIleuGlyGlnIleuSergIyGlyIySglnIySglnIySglnIySglnIySglnIy 305
Db      63 AATTTTCGGGCCAACTCTCGACTATGAGAGAAAGATTAAACACAGCTGAGACATCA 122
Qy      306 GlyProIySergIySleuIySleuHISleuGluIySProAnGluProValProAla 325
Db      123 GGGCCAAAGAGCAAACTCAAGCTGCTGACCTGAGAGCCAAAGAACCTGCTCTGCT 182
Qy      326 ValSerGluGlyGlyGlnIySergIyThrProIleuSergProProCyAlaAspSerAla 345
Db      183 GTCTCAGAGGGGTGACAGAAAGGAGAGAGCCCTCATGTCACCTGCGCAGCTGCT 242
Qy      346 ThSerGluAlaAlaGlyGlnArgProValHisProAlaSerValProSerValProSer 365
Db      243 ACCCTCAGAGGACGAGAGCAAAAGCCCGTGCATCCGCGACGCTGCCAGCTGACAGC 302
Qy      366 ValGlnProSerIleuGluIySleuSergProIleuValGlnAlaSerGlyIleuHISleu 385
Db      303 GTGCAGCGCTGCTGTTAGAGGAGCAGCCCGTGTATACAGCGCTCATGAGGGCTGCACCTG 362
Qy      386 SerAlaAspArgIleuGluIySleuSergIySleuIySArgSerPheSerIleuAspIleu 405
Db      363 TCCGCAAGACGCTGAGAGCAAGCAATTAAGCTCAAGCGTTCCTCTCTGATATCAAA 422
Qy      406 SerValSerIySergIySergIySergIySergIySergIySergIySergIySergIySerg 425
Db      423 TCAGTTTCATATTTCAGCCGAGATGCGACATCTTACATGAGCTTCTCTCATCAGAAAGAT 482
Qy      426 AlaIleuGluIyTyTyIySProSerThrIleuAspGlyThrAsnIySleuCySglnPhe 445
Db      483 GTTTGGAAATCTAACAACTTCACATCTGATGGAGCAACAAAGATTCAGAGTTTC 542

```

QY 446 SerProValGlnGluSerGlnGlnThrProGlnThrSerProAspGlnGlnAla 465
DB 543 TCCCTGTTGAGAACTATGAGACAGACTCCGAAACCACTCTGATTAAGAGAAAGCC 602
QY 466 SerIleProGlnSerGlnGlnGlnThrAlaArgProSerAspSerGlnSerIleValGlnHis 485
DB 603 AGCATCCCAAGAAAGCTGAGACAGCCGACGCTTCACAGACAGCAAGCAAGCATTCGAT 662
QY 486 SerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuSerProLeuHisArg 505
DB 663 TCGGTGAGAACAGACAGACAGTGGACACGCGCCAGAGGTCCCTTATCTCCACATGACATGA 722
QY 506 SerGlnSerValGlnAspAsnArgThrIleThrSerPhe--LeuPheGlnLeuSerThrSer 524
DB 723 AGTGGAGCCGTGAGAGCAATTAACACACACAGCTTCCTTTTGGCCCTTTCACACAGC 782
QY 525 GlnGlnHis-LeuThrIleSerSerAlaGlnLeuGlnLeuLys-GlyTrpHisSerAspIle- 543
DB 783 CAGCAGACACCTCAACGAATCTGCTGAGCTGGGAGCTTGGGCTGGGCACTTCGAATAT 842
QY 544 --LeuAlaProGlnThrSerThrPro-SerLeu-ThrSerSerTrpIlePheAlaThrGln 562
DB 843 CTTTGGGAGCCCAAGACCTTACCTTCCCTGTAACAAGAGCTGGTAATTTTGGCCACG 902
QY 562 userSerHis---PheTrpSerAlaSerAla-IleTrpGlnGln 575
DB 903 AAGCCTCACACCTTCTACTCTGCTCAGCCATTAAAGAGG 946

RESULT 5
BO933499 859 bp mRNA linear EST 21-AUG-2002
LOCUS DEFINITION AGENCOURT_8732489 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6455595
ACCESSION BO933499 5', mRNA sequence.
VERSION BO933499.1 GI:22348882
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 859)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
plate: LICM634 row: i column: 04
High quality sequence stop: 727.
Location/Qualifiers
1. 859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="6455595"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOT7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

FEATURES

source

1. 859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="6455595"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOT7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN

Alignment Scores:
Pred. No.: 8,376-110 Length: 859
Score: 1375.00 Matches: 279
Percent Similarity: 97.89% Conservative: 0
Best Local Similarity: 97.89% Mismatches: 2
Query Match: 40,234 Indels: 4
DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x BO933499 (1-859)

QY 266 SerLeuAspGlnAlaTrpArgPheValLysGlnLysArgProThrIleSerProAspPhe 285
DB 3 TCTTATGATGAGACTTACAGATTTTGTAAAGAAAGAAAGCTTACTATCTCAAACTTC 62
QY 286 AsnPheLeuGlnGlnLeuLeuAspTrpGlnLysLysIleLysAsnGlnThrGlnAlaSer 305
DB 63 AATTTCTGGGCCAATCTCCGAGCTATGAGAAAGATTAAAGAACAGACTGGAGCATCA 122
QY 306 GlnProLysSerLysLeuLysLeuLeuHisLeuGlnLysProAsnGlnProValProAla 325
DB 123 GGGCCAAAGAGCAAACTCAAGCTGTGACCTGGAGAACCAAAATGAACCTGCTCT 182
QY 326 ValSerGlnGlnGlnLysSerGlnLysProLeuSerProProGlnAlaAspSerAla 345
DB 183 GTCTCAGAGGATGAGCAAGAAAGCGAGCGCCCTTCACTCAGCTGTGCCACTTCT 242
QY 346 ThrSerGlnAlaAlaGlnArgProValHisProAlaSerValProSerValProSer 365
DB 243 ACCTCAGAGGAGCAGAGCAAGAAAGCCGTGACATCCGCCAGGCTGCCAGCGCCAGC 302
QY 366 ValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeuSerGlnLysLeu 385
DB 303 GTGCACCGCTGCTGTGAAGAGCACCCGCTGTGACAGCGCTCAGTGGGCTGCACCTG 362
QY 386 SerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPheSerLeuAspIleLys 405
DB 363 TCCGACAGAGGCTGAGAAAGACAAATAGCTCAACGCTTCTTCTCTGATATCAAA 422
QY 406 SerValSerTrpSerAlaSerMetAlaAlaSerLeuHisGlnPheSerSerSerGlnAsp 425
DB 423 TCACTTCAATATTCACCGACGACATGCGACATCTTCAATGAGCTTCATCAGAAAT 482
QY 426 AlaLeuGlnTrpTrpLysProSerThrThrLeuAspGlnThrAsnLysLeuGlnPhe 445
DB 483 GTTTTGAATTAATAACAACCTTCACTACTGTGATGGAGCAACAAGCTATGCCAGTTC 542
QY 446 SerProValGlnGluSerGlnGlnThrProGlnThrSerProAspGlnGlnAla 465
DB 543 TCCCTGTTGAGAACTATGAGACAGACTCCGAAACCACTCTGATTAAGAGAAAGCC 602
QY 466 SerIleProGlnSerGlnGlnGlnThrAlaArgProSerAspSerGlnSerIleValGlnHis 485
DB 603 AGCATCCCAAGAAAGCTGAGACAGCCGACGCTTCACAGACAGCAAGCAAGCATTCGAT 662
QY 486 SerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuSerProLeuHisArg 505
DB 663 TCGGTGAGAACAGACAGACAGTGGACACGCGCCAGAGGTCCCTTATCTCCACATGACATGA 722
QY 506 SerGlnSerValGlnAspAsnArgThrIleThrSerPheGlnLeuSerThrSerGln 525
DB 723 AGTGGAGCCGTGAGAGCAATTAACACACACAGCTTCTTTTGGCCCTTTCACACAGC 782
QY 526 GlnHisLeuThrIleSerSerAlaGlnLeuGlnLeuLys-GlyTrpHisSerAspIle 544
DB 783 CAGCAGACACAGACTCTGCTGAGCTTGGGAGCTTAAAGGCTGAGNACCTCGATATTC 842
QY 544 ValaPro 546
DB 843 TGGGCCC 849

RESULT 6

```

row: 1 column: 08
High quality sequence stop: 669.
Location/Qualifiers
1. 898

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6194455"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="DN10B"
/lab_host="DH10B"
/clone_lib="lupsk4; sympathetic trunk"
/note="vector: pCMV-SPOK2 (Life Technologies); Site_1:
North; Site 2: Salt; cDNA made by oligo-dT priming.
directionally cloned using the following adaptors:
5'-TCGACCCACGGGCTCG-3' and
5'-GACGTGTTTCATGATCGGACGCGCCGCCCT(15)-3'. Site selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

```

Alignment Scores:	
Pred. No.:	2.74e-109
Score:	1369.50
Percent Similarity:	95.27%
Best Local Similarity:	94.26%
Query Match:	40.07%
DB:	13
	Gap
	4
	12
	Mismatches:
	Conservative:
	Matches:
	Length:
	898

88 10-025-345A-109 (1-665) x BQ721265 (1-898)

QY 4 G I U N E C I I E G I Y T H R G I I I L E V A L T h G I A R g L e u V a I A L a L e u e n g I S e r G I y T h R 23
Db 14 G A G A T G A T T G A A C T C A A T T G T T A C G A G A G G T G G T G C T T G C T G A A A G T G A A C G 73
QY 24 G I U I y S V a I L e u L e u I I e a S S e r A n g P r o P h e V a I G I U r y A S n T h S e r S i I L e u 43
Db 74 G A A A A T G C T G C T A A T T G A T T A C C G C C A T T T G A A T A C A A T A C A T C C A C A T T T T G 133
QY 44 G I U A I A I L e a S n I L e a S C y S e r I y S L e u M e I y A n g A n g L e u G I n G I n A S p I y S V a I 63
Db 134 G A A G C C A T T A T A T C A A C T G C T C A A C G C T T A G A A C G A A G G T T C C A A C A G A C A A A G T G 192
QY 64 L e u I I e T h G I L e u I I e G I n H I S e r a l a y S h I S y V a I A S p I I e A S P C y S e r G I n 83

Dp	194	TTAATTACAGAGCTATCCAGCATTCAAGCGAAATAAGTTGACATTGATTCGACGTAC	253
Oy	84	LysValIvalValTyrAspLnsSerSerGlnaspValAlaserLeusSerSeraspCySphe	103
Dp	254	AAGGTGGTGATTGAATGCATCAAAGCTCCCAAAATGTTCCTCTCTCTTCAAGACTGTTTT	313
Oy	104	LeuThrValLeuLeuGlyLyseLeuGluLylserPheasnSerValHisLeuLeuAlagly	123
Dp	314	CTCACGTGATCTTCTGGGTAACAAGGCTTCAACTCTGTCAACCTGACCTGCCTTCAGGT	373
Oy	124	GlyPheAlaGlupheserargCySpheProglyLeucSeglueGlyLysSerthrLeuVal	143
Dp	374	GGGTTTGCTGAGTTCTCTCGTTGTTTCTCGGCTCTGTGAAGAAGAAATCCACTGATGC	433
Oy	144	ProthrCySllSerEgInProCySleuProvalAlAasnllleglProthrarglileu	163
Dp	434	CTTACCTCATTTCTCGACCTTGCTTACCCTGTGTGCCAACATGGGGCAACCCGAATTCCT	493
Oy	164	ProasneuryrleugilyCySglnarGaSPValLeuasnlYsgluleuIllednglnasn	183
Dp	494	CCCAATCTTATCTTGCTGGCCGACGAGATGTCTTCAACAAGACCTGATGCAGCAAT	553
Oy	184	GlyTlleglyTyralLeuasnalaserTythrCySprolysProaspPheIIeproglu	203
Dp	554	GGGATTTGGTATGTGTTAATGGCAGCAATACCTTCTCAAAAGCTGCATTTATCCCCGAG	613
Oy	204	SerHisPheLeuharyalProvalasnpsSerPhecYsgluYslleLeuProtpPleu	223
Dp	614	TTCATTTCTCGCGGTGCTGTGAATACAGCTTTGTGAGAAATT-TTGGCGTGGTTG	672
Oy	224	AspLysSerValaspPheIleGluLYsaIalaLYsaIaseRanglyCySValleuValHis	243
Dp	673	GACAAATCGAGTATTCATTGAGAAACAAAACCTCCANTGASUTGTGTTAAGTGCAC	732
Oy	244	CYSLeuAlagIlyIseserSarSerAlahrrllealleallealyrllemelysarGmet	263
Dp	733	TGTTTAGCTGGGATCTCCCCTCGGC-ACCAATGCTATCCCTACATCATGGAAGAGATG	791
Oy	264	AspMetSer-LeuaspGluAlaryrarGpheValylYsgluLYsarGProthrIIsesPr	283
Dp	792	GACATGCTCTTTTAATGAACCTACGATTTGTGGAAAAAGAAAAGCCTACATATTC	851
Oy	283	GasnpPheasn--PheleuglyGlnleuLeuaspTyrgluYslYlsYs	297
Dp	852	AAAACTCAATTTTCTTGGGCCAACTCCTGGCATTAATGAAAAA	897
RESULT 7			
BG482429			
LOCUS	BG482429	903 bp	mRNA linear EST 21-MAR-2001
DEFINITION	605285595F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4650430 5'		
ACCESSION	BG482429		
VERSION	BG482429.1	GI:13414708	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgi.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
CONTACT:	Robert Strausberg, Ph.D		

cDNA library preparation: Ling Hong/Rubin Laboratory
DNA sequencing by: The I.M.A.G.E. Consortium (LINTL)
Clone distribution: Incyte Genomics, Inc.
Found through the I.M.A.G.E. Consortium information can be
http://image.llnl.gov
Plate: LINC431 row: m column: 23
High quality sequence stop: 806.

FEATURES
source

Location/Qualifiers
1. 903
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4650430"
/issue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:

Pred. No.: 1.88e-108 Length: 903
Score: 1360.00 Matches: 281
Percent Similarity: 94.72% Conservative: 6
Best Local Similarity: 92.74% Mismatches: 10
Query Match: 39.79% Indels: 6
DB: 12 Gaps: 2

US-10-029-345a-109 (1-665) x BG482429 (1-903)

Qy 172 ArgAspValLeuAsnLysGluLeuIleGlnGlnAsnGlyIleGlyIleValLeuAsnAla 191
Db 3 CGAGATCTCTCAACAGAGAGCTGATCCAGCAAGATGGATTGTTATGTGTTAAATGCC 62
Qy 192 SerTyrThrCysProLysProAspPheIleProGlnSerHisPheLeuArgValProVal 211
Db 63 AGCAATACCTGTCCAAAGCCTGACTTATCCCGAGTCTCATTTCTCGCTGTGCTGTG 122
Qy 212 AsnAspSerPheCysGluLysIleLeuProTyrLeuAspLysSerValAspPheIleGlu 231
Db 123 AATGACACGCTTTGTGAGAAATTTTGCCTGTGTGAGCAAAATCAGTAGATTTTCATTGAG 182
Qy 232 LysAlaLysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSer 251
Db 183 AAAGCAAAAGCCTCCATGATGATGTTCTAGTGCACTGTTAGCTGGATCTCCCGCTCC 242
Qy 252 AlaThrIleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyr 271
Db 243 GCCACCATCGCTATCGCTCATCATCATGAGATGAGATGATGCTTTAGATGAAGCTTAC 302
Qy 272 ArgPheValLysGluLysArgProThrIleSerProAsnAspPheLeuGlyGlnLeu 291
Db 303 AGATTGTGAAAGAAAAGAACCTCACTATCTCCAAACCTCAATTTCTGGGCGCACTC 362
Qy 292 LeuAspTyrGluLysLysIleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeu 311
Db 363 CTGGACCTATAGAGAAAGATTAGAACCAAGCTGAGCATCAGGCGCAAGAGCAAACTC 422
Qy 312 LysLeuLeuHisLeuGluLysProAsnGluProValProAlaValSerGluGlyGln 331
Db 423 AAGCTGTGCACTCGAGAAAGCAAAATGAACCTGCTCTGCTCTCAGAGGGGTGACAG 482
Qy 332 LysSerGluThrProLysSerProProCysAlaAspSerAlaThrSerGluAlaAlaGly 351
Db 483 AAAAGCAGAGCGCCCTCAGTCCACCTGTGCCAAGCTCTCTCTCAGAGGAGCAGCGGA 542
Qy 352 GluArgProValHisProAlaSerValProSerValProSerValGlnProSerLeuLeu 371
Db 543 CAAGAGCCCTGCTATCCCGCAGGCTGCCAGGCTGCCAGGCTGCGCTGCTTTA 602
Qy 372 GluAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGlu 391
Db 603 GAGGACAGCCCGCTGTACAGGCGCTCAGTGGGCTGACCTGTCCGAGACAGGCTGAA 662

Qy 392 AspSerAsnLysLeuLysArgSerPheSerLeuAspIleLysSerValSerTyrSerAla 411
Db 663 GACACAGATAGCTCAAGGCTTCTCTCTCGAGATCAATCCAGTTTCATATTCAGC 722
Qy 411 AsnMetAlaAlaSerLeuHis---GlyPheSerSerSerGluAsp-AlaLeuGluTyrT 430
Db 723 CAGCATGGACAGATCCCTTACATGTGCTTCATCTCATCAGAAATGCTTTGGAATACT 782
Qy 430 YrlAspPro-SerThr-ThrLeuAsp---GlyThrAsnLysLeuCysGlnPheSerProVa 448
Db 783 ACAAACTTTCACCTTACTCTGATTCGACCAAGCAAGATTATGTCAGTTCTCCCTGT 842
Qy 448 LglnGluLysSerGluGlnThrProGlnThrSerProAspLysGluGluAlaSerIlePr 468
Db 843 TCGAGACTATTCGAGAAAGACTCCGAAACCACTCTGTATAGGAGAGAAAGCATATCCC 902
Qy 468 o 468
Db 903 A 903

RESULT 8
BC038231
LOCUS
DEFINITION
Homo sapiens, similar to dual specificity phosphatase 8, clone
IMAGE:5547764, mRNA.
ACCESSION
BC038231.1 GI:23398534
VERSION
BC038231.1 GI:23398534
KEYWORDS
HTC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2207)
Strausberg, R.
Direct Submision
Submitted (30-SEP-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgti.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaepi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W.,
Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

REMARK
COMMENT
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRK Plate: 79 Row: a Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4758211
This clone has the following problem: frame shifted.

FEATURES
source

Location/Qualifiers
1. 2207
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5547764"
/issue_type="Skin, melanotic melanoma."
/clone_id="NH_MGC_72"

Qy 630 MetGluPheGlySerIleMetSerGluAsnArgSerArg---GluGluLeuGlyLys 648
 Db 1899 ATGAGAGTTCCAGAGAGGC---ATGCTGAGAGGCGCGCGCGCGAGAGCTGGCGCC 1955
 Qy 649 ValGlySerGlnSerPheSerGlySerMetGluIleIleGluValSer 665
 Db 1956 CTGGGCGAGAGGAGGAGCTTCTCGGCGAGCGTGAAGCTCATCGAGGTCTCC 2006

RESULT 9

BQ770036

LOCUS BQ770036 862 bp mRNA linear EST 26-JUL-2002
 DEFINITION UI-M-FIO-bz-o-24-0-UI.r1 NIH BMAP_FIO Mus musculus CDNA clone
 IMAGE:5702255 5', mRNA sequence.

ACCESSION

BQ770036

VERSION

BQ770036.1

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

location/Qualifiers

1.862

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5702255"

/tissue_type="whole brain"

/dev_stage="embryo 12.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_1lb="NIH BMAP_FIO"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bontade, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCAGAGC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

US-10-029-345A-109 (1-665) x BQ770036 (1-862)

Qy 116 AsnSerValHisIleuLeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeu 135
 Db 8 CACTGTGTCACCTGCTGTGAGGAGGCTTGTGAGAGTCTCTGTGTTTCCCTGGGCTC 67
 Qy 136 CysGluGlyLysSerThrIleuValProThrCysIleSerGlnProCysLeuProValAla 155
 Db 68 TGTGAAGAAAGTCCACTGCTAGTCCCTACCTGATATCTCAGCTTCTTACCTGTTGGC 127

RESULT 10

BUT04078

LOCUS BUT04078 769 bp mRNA linear EST 15-JUL-2003
 DEFINITION UI-M-FIO-bz-n-23-0-UI.r1 NIH BMAP_FIO Mus musculus CDNA clone
 IMAGE:6406486 5', mRNA sequence.

ACCESSION

BUT04078

VERSION

BUT04078.1

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

location/Qualifiers

1.862

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5702255"

/tissue_type="whole brain"

/dev_stage="embryo 12.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_1lb="NIH BMAP_FIO"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bontade, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCAGAGC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

US-10-029-345A-109 (1-665) x BUT04078 (1-862)

REFERENCE 1 (bases 1 to 769)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLM at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES
 Source Seq primer: pYX-5.
 Location/Qualifiers

1..769
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6406486"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1lb="NIH-BMAP-F00"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: Scov I;
 Site 2: Not I; The library was constructed according
 to Bontado, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with BclI adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is TGAGAGAGCC. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemlin Chan, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 6,06e-96 Length: 769
 Score: 1216.00 Matches: 229
 Percent Similarity: 94.09% Conservative: 10
 Best Local Similarity: 90.16% Mismatches: 15
 Query Match: 35.58% Indels: 0
 DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x B0704078 (1-769)

QY 93 GlnaPValAlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGlyValLeuGlu 112
 DB 6 CAAGATGTTGGTCTCTCTCTGACAGCTGCTTCTCACTGACTTCTGGGTAACCTGAG 65
 QY 113 LysSerPheAaSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerAaGysPhe 132
 DB 66 AGGAGCTTCAACTGCTGCTCACTGCTGAGGTGGCTTCTGAGTTCTCTGTTTTC 125
 QY 133 ProGlyLeuGysGluGlyLysSerThrLeuValProThrCysHisSerGlnProCysLeu 152
 DB 126 CTTGGCTCTGTGAAGAGAAAGTCACTTGAAGTCCCACTGCAATCTGAGCTTGA 185
 QY 153 ProValAlaAaenilegylProThrArgileLeuProAaenileuGlyCysGlnArg 172
 DB 186 CCGTGTTCGAACATTTGGGCACTCGAATCTTCCCAATCTCATATTTGGCTGCAAGCA 245
 QY 173 AspValLeuAaenilegylLeuGlnAaenilegylTyrValLeuAaenileuAaSer 192
 DB 246 GATGTCTTCAACAGAGCTGATGCAACAGAAATGGATGGCTATGTGTTAATCCAGC 305

QY 193 TyrThrCysProLysProAspPheIleProGluSerHisPheLeuArgValProValAsn 212
 DB 306 AATACCTGTCCAAAGAGCCGACTTATCTATCTCAATCTCACTCCGCAAGGCTGGAAT 365
 QY 213 AppSerPheCysGluValIleLeuProTyrLeuAaPysSerValAppPheileGlyLys 232
 DB 366 GACACCTTTTGTGAAGAAATCTTCAACATGCTGTGGAACAAGTCTGAGATTTCATTGAGAA 425
 QY 233 AlalysAlaSerAaenilegylCysValLeuValHisCysLeuAlaGlyLysSerAaSerAla 252
 DB 426 GCAAAAGCTTCATATGCTGTGCTGCTTATCACTGCTTACTGCTGAGATCTTCTGCTCCGCC 485
 QY 253 ThrIleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAaPysGluAlaTyrArg 272
 DB 486 ACTATGTCTATTTGGCTTCACTATCAATGAAGAGATGAGCATGTCTTATATAGCTTACAGA 545
 QY 273 PheValLysGluLysAaProThrIleSerProAaPheAaPheLeuGlyGlnLeuLeu 292
 DB 546 TTTGTGAAGAAAGAAAGAACTTACTATATCTCCGAATTTAATTTATGCGCAACTCATG 605
 QY 293 AspTyrGluLysLysIleLysAaenileuThrGlyAlaSerGlyProLysSerLysLys 312
 DB 606 GACTATGAGAAAGCAATTATATACCACTGGAATGTCAGGCCCAAGAGCAACTGAGAG 665
 QY 313 LeuLeuHisLeuLeuLysProAaenileuProValProAlaValSerGluGlyGlyLys 332
 DB 666 CTCTGCACTTACCAACCACTGAGTGAAGCCGCTGCAAGCTTCAAGAGGCGGATGAGAG 725
 QY 333 SerGluThrProLeuSerProProCysAlaAaPysSerAlaThr 346
 DB 726 AGTGAATGTCTCTGCTGCTCACTCTGAGCACTGAGCACTGAGCTG 767

RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 836)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLM at:
<http://image.llnl.gov>
 Plate: LLM1440 row: f column: 21
 High quality sequence stop: 805.
 Location/Qualifiers

FEATURES

source

1..836
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5176724"
 /lab_host="DH10B"
 /clone_1lb="NIH-MGC 115"
 /note="Organ: Pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is

destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	1e-92	Length:	836
Score:	1180.00	Matches:	260
Percent Similarity:	93.97%	Conservative:	5
Best Local Similarity:	92.20%	Mismatches:	15
Query Match:	34.52%	Indels:	9
DB:	12	Gaps:	0

US-10-029-345A-109 (1-665) x B1821804 (1-836)

```

Qy 50 CysSerIysLeuMetLysArgArgLeuGlnGlnAspIysValIleThrGluLeuIle 69
Db 2 TGCTCCAGCTTATGAGCGAGAGGTTCCACAGACAAAGCTTATTCAGAGCTCATC 61
Qy 70 GlnHisSerAlaIysHisIysValAspIleAspCysSerGlnIysValIValIYrAsp 89
Db 62 CAGCATTCAGCGAAACATTAAGTTGACATTGA-TGCAGTCAGAAAGCTTGTAGTTACGAT 120
Qy 90 GlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrValIleLeuGly 109
Db 121 CAAAGCTCCCAAGATGT-GCCTCTCTCTTCACAGATGTTTCTCAGCTGACTTCGTGGT 179
Qy 110 LysLeuGlnIysSerPheAsnSerValHisIleLeuValIagIyIlePheAlaGluPheSer 129
Db 180 AAACGTGAGAAAGCTTCAACTCTGTTCACTGCTTCAGAGGCTTTCAGATTCTCT 239
Qy 130 ArgCysPheProGlyLeuGlyGlyIysSerThrLeuValProThrCysIleSerGln 149
Db 240 CGTAGTTTCCCTGGCCTCTGTGAAGAAATCCACTAGCTCCTACCTGCACTTCTCAG 259
Qy 150 ProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuIYrLeuGly 169
Db 300 CTTGTCTACCTGTGTGCAACATTTGGGCAACCCGAAATTTCTCCAAATCTTTATCTGGC 359
Qy 170 CysGlnArgAspValIleAsnIysGluLeuIleGlnIleAsnGlyIleGlyIYrValIleu 189
Db 360 TGCCAGCGAATGCTCTCAACAGAGAGCTGATGACGAAATGGATCGGTATGTCTTA 419
Qy 190 AsnAlaSerIYrThrCysProIysPProAspPheIleProGluSerHisPheLeuArgVal 209
Db 420 AATGCCAGCAATACCTGTCCAAAGCTGACTTATCCCGAGTCTCATTTCTGCGTGTG 479
Qy 210 ProValAsnAspSerPheCysGluIysIleLeuProThrIleLeuAspIysSerValAspPhe 229
Db 480 CCGTGTGAATACAGCTTCTGTGAGAAATTTGGCGGTGGCAAAATCAGATGATTTTC 539
Qy 230 IleGluIysAlaIysAlaSerAsnGlyCysValIleValHisCysLeuAlaGlyIleSer 249
Db 540 ATTGAGAAAGCAAAAGCTTCAATGATGTGT-CTAAGTGAAGCTTTAGCTGGGATCTCC 598
Qy 250 ArgSerAlaThrIleAlaIleAlaIYrIleLeuIysArgMetAspMetSerLeuAspGlu 269
Db 599 CGCTCCGC-ACCATCGCTATCGC-TACATCATGAAAGAGATGACATGCTTTAGATGAA 656
Qy 270 AlaIYrArgPheValIysGlu-LysArgProThrIleSerProAsnPheAsnPheLeu-G 289
Db 657 GCTTACGATTTTGGACAGAACAGACCTACATATCTCCCAAACTTCAATTTTCTGGG 716
Qy 289 LysIleLeuLeuAspIYrGluIysIysIleIysAsnGlnThrGlyAlaSerGlyProIys 309
Db 717 GCCAGCTCTCTGAGACTATGAGAGAAATTAAAGAACGAGCTGGAGCATCAGGGGC-AGA 775
Qy 309 erIysLeuIysLeuHisIleGlnIulysProAsnGluProValProAlaValSerGlu 329
Db 776 GCTTAATCAAGCTCTGCTCAGCTGGG-AGAAAGCATATGAACTGCTGCTGTCTCAAGG 834

```

Qy 329 IY 329
Db 835 GG 836

RESULT 12

LOCUS

DEFINITION

CF727177 715 bp mRNA linear EST 09-OCT-2003
 UI-M-HBO-ck-j-09-0-UT-r1 NIH_BMAP_HBO Mus musculus cDNA clone
 IMAGE:30548096 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Seq primer: pYX-5.

Location/Qualifiers

1. 715

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:30548096"

/tissue_type="whole eye"

/dev_stage="embryo 12.5,13.5,14.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP HBO"

/note="Organ: Eye; Vector: pYX-Asc; Site 1: Bcor I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Bcor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TATATGAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Alignment Scores:

Pred. No.:	4.91e-90	Length:	715
Score:	1148.00 <td>Matches:</td> <td>223</td>	Matches:	223
Percent Similarity:	96.64%	Conservative:	7
Best Local Similarity:	93.70%	Mismatches:	8
Query Match:	33.59%	Indels:	2
DB:	14	Gaps:	0

US-10-029-345A-109 (1-665) x CF727177 (1-715)

```

Qy 87 ValIYrAspGlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrVal 106
Db 2 GTTATGATCAAAAGTCCCAAGATGTGTTCTCTCGCAGACAGCTTTCACACTGTA 61
Qy 107 LeuLeuGlyIysSerPheAsnSerValHisIleLeuValIagIyIlePheAla 126

```

```
Db      62 CTTCTGGGTAACAGGAGAGGCTTCAACTGCTGACGCTTGACGGGCTTGTGCT 121
Qy      127 GlnPheSerArgCysPheProGlyLeuGlyLeuGlySerThrLeuValProThrCys 146
Db      122 GAGTCTCTGCTGTTTCTGCTGCTGTAAGGAAAGTCACTAGTCCCTCACTGCG 181
Qy      147 ILeSerGlnProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeu 166
Db      182 ATATCTGCGCTTGTACCTGTTGCGAACAATGGGCGCAACTCGAATCTTCCCAATTC 241
Qy      167 TrrLeuGlyCysGlnArgAspValIleuAsnIleuGlyIleGlnIleAsnGlyIleGly 186
Db      242 TATCTGGCTCCGACGAGATGCTCTCAACAAGGACCTGATGACAAAGAAATGGGATGGCG 301
Qy      187 TyrValIleuAsnAlaSerTrrThrCysProIlyProAspPheIleProGlnSerHisPhe 206
Db      302 TATGTTAAATGCCAGAAATPACCTGTCCAAAGCCTGCAATCTTCACTGAAATCTCACTTC 361
Qy      207 LeuArgValProValAsnAspSerPheCysGlyIleuValIleuProTrrLeuAspIlySer 226
Db      362 CTGGAGAGCTCCTGTGAATGACAGCTTTTGTGAGAAATCCTACATGGTTGGACAAAGTCT 421
Qy      227 ValAspPheIleGlyIleValAlaIleAsnGlyCysValIleuValHisCysIleuAla 246
Db      422 GTGCAATTCATTTGAGAAAGCAAAAGCCTCAATGGCTGTGTATCTCACTGCTTACGCT 481
Qy      247 GlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaTrrIleMetIlyArgMetAspMetSer 266
Db      482 GGGATCTCGCTCGCCACATTTGCTATGCTTACATCAATGAAAGAGATGGAACATGCT 541
Qy      267 LeuAspGlnAlaTyrArgPheValIleGlyIleValAspProThrIleSerProAsnPheAsn 286
Db      542 CTAGATGAGCTTACAGATTGTGAAAGAAAAAGACCTACTATATCTCCGAATTTTAT 601
Qy      287 PheLeuGlyGlnIleuLeuAspTrrGlyIleValIleValAsnGlnIleThrGlyAlaSerGly 306
Db      602 TTTATGGGCAACCACTAGACTATGAGAAAGCATTAATACCAAGACTGGAATGCTCAGGG 661
Qy      307 ProIlySerIlyLeuIlyLeuLeuHisIleuGlyIleuProAsnGlnProValPro 324
Db      662 CCANAGAGCAA-CTGAACTGCTGCACTTACAGAACCAAGT-GAGCCCGTGCT 713

RESULT 13
LOCUS   BE897795 920 bp mRNA linear EST 20-OCT-2000
DEFINITION M891438457F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3923247 5',
ACCESSION BE897795
VERSION BE897795.1 GI:10363618
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 920)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
          Email: gsa@bbs-rcmail.nih.gov
          Tissue Procurement: ATCC/DCTP/DRP
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: LHAM9759 row: b column: 16
          High quality sequence stop: 678.
          Location/Qualifiers
            1..920
            /organism="Homo sapiens"
```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3923247"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
```

ORIGIN

Alignment Scores:

Pred. No.:	7,36e-86	Length:	920
Score:	1102.50	Matches:	235
Percent Similarity:	89.30%	Conservative:	7
Best Local Similarity:	86.72%	Mismatches:	19
Query Match:	32.26%	Indels:	11
DB:	10	Gaps:	3

US-10-029-345A-109 (1-665) x BE897795 (1-920)

```
Qy      233 AlAluYAlaser-AsnGlyCysValIleuValHisCysIleuAlaGlyIleSerArgSerAl 252
Db      2  GCAAAAGCCCTCCGGTGGATGTGTTCTAGTGAAGCTGTTAGCTGGGATCTCCGCTCCGC 61
Qy      252 aThrIleAlaIleAlaTrrIleMetIlyArgMetAspMetSerIleuAspGlnAlaTyrArg 272
Db      62 CACCATCGCTATGCTTACATCAATGAAAGAGATGACATGCTTTAGATGAACCTTACAG 121
Qy      272 gPheValIlyGlyIlyValArgProThrIleSerProAsnPheAsnIleuGlyIleuLeu 292
Db      122 ATTGTGAAGAAAGAAAGAACCTACTATATCTCAAACTTCAATTTCTGGGCCAACCTCT 181
Qy      292 uAspTrrGlyIlyValIleValAsnGlnIleThrGlyAlaSerGlyProIlySerIlyLeu 312
Db      182 GGACTAGAAAGAAAGATTGAAGCAAGCATGAGATCAAGGCGCAAGAGCAAACTCAA 241
Qy      312 sIleuLeuHisIleuGlyIlyValProAsnGlnProValProAlaValSerGlyGlyGlnIly 332
Db      242 GCTGCTGCACCTGGAAGAGCAATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
Qy      332 sSerGlnThrProLeuSerProProCysAlaAspSerAlaThrSerGlnAlaIleGly 352
Db      302 AAGCGAGAGCGCCCTCAGTCCACCTGTGCGGCACTGCTACTCACTCAGAGGACAGAGCA 361
Qy      352 uArgProValHisProAlaSerValProSerValProSerValGlnProSerIleuGly 372
Db      362 AAGGCCCGTGCATCCGCC-----AGCGTGGCCAGCGTGCAGCCGCTGTTAGA 412
Qy      372 uAspSerProLeuValAlaIleuSerGlyLeuHisIleuSerAlaAspArgIleuGly 392
Db      413 GACAGCGCGCTGTTACAGCGCTCAGTGGCGTGCAGCTGTGCGAGACAGGCTGGAAGA 472
Qy      392 pSerAsnIlyLeuIlyValArgSerPheSerIleuAspIleValSerValSerTyrSerIase 412
Db      473 CACCAATTAAGCTCAACAGTTCCTCTCTGATTAACAATCAGTTTCATATTACACCCG 532
Qy      412 tMetAlaAlaSerLeuHisGlyPheSerSerSerGlnuAspAlaIleuGlyTrrTrrTrrPr 432
Db      533 CATTGGAGCAATCTTACATAGCTTCTCTCATCAGAAAGAGCTTGGAAATACACAAACC 592
Qy      432 oSerThrThrLeuAspGlyThrAsnIlyLeuGlyCysGlnPheSerProValGlnIleuSe 452
Db      593 TTCCACTACTCTGAGAGGAGCAACAGCTATGCCGTTCTC-CTGTCCAGGAATATAC 651
Qy      452 rGlnGlnThrProGlnThrSerTrrAspIlyGlyIleuAlaSerIleProIlyIlyLeuGly 472
Db      652 GGAGCAGACTCCCGAA---CAAGTCAATAGAGGAGCAACAGCATCCCAAGAGGCTGGA 708
Qy      472 nThrAlaArgProSerAspSerGlnSerIlyValIleuHisSerValArgThrSerSerSe 492
Db      709 GACCGAGGGGCTTACAGAAAGCCAGA-----GCAGCGATGCTTTCGCTCAA 753
```

```

Oy      492 rGlyThrValaglnArgserLeuLeuSerPro 502
          |||
          :|:::|||||
Db      754 CCCACAGGTGGCACGCCACGAGTCTTTCCTCC 784

RESULT 14
LOCUS      B1816954/c
DEFINITION  imgseqc_10_2000/g1z410bdf1.xl Soares_NPBMC Homo sapiens cDNA
ACCESSION  B1816954
VERSION    B1816954.1
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS    Kale,P.I., Harsch,T.J., Folta,P.A., Nelson,D.O., Sanders,C.G. and
           Prange,C.K.
TITLE      The I.M.A.G.E. Consortium quality control effort: clone
           reresequencing for verification
JOURNAL    Unpublished (2001)
COMMENT    Other ESTs: BG058779
           Contact: Prange CK
           The I.M.A.G.E. Consortium
           Lawrence Livermore National Laboratory
           Livermore, CA, USA
           Email: help@image.llnl.gov
           This read has been verified (found to hit its original self in the
           correct orientation), as part of the I.M.A.G.E. Consortium quality
           control effort. High quality sequence is defined as having 100 or
           more base pairs with a phred quality value of 20 or greater, where
           a sliding window of 4 base pairs with a phred quality value of 15
           or greater marks the beginning and end of the sequence. For
           information on obtaining this clone, please contact
           info@image.llnl.gov. effort.
           Plate: LLM9388 row: k column: 7
           Seq primer: -21m13
           High quality sequence stop: 682.
           Location/Qualifiers
             1..682
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:4140798"
               /tissue_type="lymphocyte"
               /lab_host="DH10B (phage-resistant)"
               /clone_1lb="Soares_NPBMC"
               /note="Organ: blood; Vector: pUT73D-Pac; Site 1: NotI;
               Site 2: EcoRI; let strand cDNA was primed with a Not I -
               oligo(dT) primer [5',
               TGTTACCAATCTGAAGTGGAGCGCGCGGTTTCTTTTCTTTTCTTTTCTTTTCTTTT
               3']; double-stranded cDNA was ligated to Eco RI adaptors
               (Pharmacia), digested with Not I and cloned into the Not
               I and Eco RI sites of the modified pUT73 vector. library
               is normalized; constructed in the laboratory of M. Bento
               Soares (University of Iowa)."
```

[illegible]

source

```

1. .656
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30357529"
/tissue_type="whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP GH0"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAAGTGAAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

```

ORIGIN

Alignment Scores:

```

Pred. No.:      3,18e-84      Length:      656
Score:          1081.00      Matches:      208
Percent Similarity: 98.60%      Conservative: 4
Best Local Similarity: 96.74%      Mismatches: 2
Query Match:      31.63%      Indels:      1
DB:              14          Gaps:      0

```

US-10-029-345A-109 (1-665) x CF532917 (1-656)

```

QY      16 ValAlaLeuLeuGluSerGlyThrGluValLeuLeuLeuLeuAspSerArgPro-PheVal 35
Db      3 GTGGCTCTGCTGGAAGTGAACGAAAAAGTCTGTAATTGATGAGCCGACATTGTTGT 62
QY      35 IGlutyrAsnThrSerHisIleLeuGluAlaIleAsnIleAsnCySerIysIleuMetIly 55
Db      63 GGAATACAAATACGTCCTACATTTGGAAAGCCATTAAATACACGCTCCAAACGATGAA 122
QY      55 sArgArgLeuGlnGlnAspIysValLeuIleThrGluLeuLeuLeuGlnHisSerAlaIysHis 75
Db      123 GCGAAGGTGCAACGACGAAAGATTTATACAGAACTTATCCAACTTCTGCAAAAGCA 182
QY      75 sLysValAspIleAspCySerGlnIysValValValItyrAspGlnSerSerGlnAspVal 95
Db      183 TAAAGTTGACATTCATGTCATCAGAAAGTGTGATGATGATCAAGTCCCAAGATGT 242
QY      95 lAlaSerLeuSerSerAspCySphenLeuThrValLeuLeuGlyIysLeuGluIysSerPh 115
Db      243 TGGTTCCTGTCGTCAGCTGCTTCTCCTACCTGTAACCTTGGTAAACCTGAGAGAGACTT 302
QY      115 eAsnSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArgCySphenProGlyLe 135
Db      303 CAATCTGTGTCACCTGCTGAGAGGTGGTTGCTGAGTTCCTCTGTTTCCCTGGGCT 362
QY      135 uCyGluGlyIysSerThrLeuValProThrCysIleSerGlnProCySerLeuProValAl 155
Db      363 CTGTGAAGAGAAATCCACTAGTCCCTACCTGCATATCTCAGGCTTGCTTACCTGTTGC 422
QY      155 aAsnIleGlyProThrArgIleLeuProAsnLeuThrLeuGlyCySglnArgAspValIle 175
Db      423 GAACATTGGGCAACTGAAATCTTCCCAATCTCTATCTGGCTGCCAGCGAGATGTCCT 482
QY      175 uAsnIysGluLeuLeuIleGlnIleAsnGlyIleGlyTyrValIleuAsnAlaSerTyrThrCy 195
Db      483 CAACAAAGAGACTGATGCAACGAAATGGGATTGGCTATGTGTAAATGCCAGCAATACCTG 542

```

```

QY      195 sProIysProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPh 215
Db      543 TCCAAGCTGACTTACCTGATCTCTCTGCGAGTGCCTGTGAATGACAGCTT 602
QY      215 eCySglnIysIleLeuProThrIleuAspIysSerValAspPhe 229
Db      603 TTGTGAGAAAATCCTACCATGTGTGACAAAGTCTGTGGATTTC 645

```

Search completed: June 21, 2004, 23:33:27
Job time : 4997.03 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004, Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:19:48 ; Search time 57.0786 Seconds

(Without alignments)
3289.124 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 3418
Sequence: 1 MAHEMIGTQVTRVALLE.....LKVSGSSRSGSMELIEVS 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppa/PCR_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppa/PCRUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	3406	99.6	665	9 US-09-816-494-2	Sequence 2, Appl1
2	3406	99.6	665	9 US-09-964-277-2	Sequence 2, Appl1
3	3406	99.6	665	12 US-10-072-012-680	Sequence 680, App
4	3406	99.6	665	12 US-10-168-506-14	Sequence 14, Appl
5	3406	99.6	665	12 US-10-343-357-7	Sequence 26, Appl
6	3406	99.6	665	15 US-10-377-072-26	Sequence 26, Appl
7	3406	99.6	665	16 US-10-257-026-2	Sequence 2, Appl1
8	3406	99.6	665	16 US-10-648-593-240	Sequence 240, App
9	3406	99.6	665	12 US-10-648-593-247	Sequence 247, App
10	3406	99.6	665	12 US-10-072-012-679	Sequence 679, App
11	3406	99.6	665	12 US-10-072-012-703	Sequence 703, App
12	3406	99.6	665	12 US-10-425-114-54204	Sequence 54204, A
13	3399	99.4	665	12 US-10-072-012-681	Sequence 681, App
14	3399	99.4	665	15 US-10-094-749-2312	Sequence 2312, App
15	3379.5	98.9	662	12 US-10-072-012-258	Sequence 258, App

16	3358.5	98.3	680	12 US-10-072-012-256	Sequence 256, App
17	3079.5	90.1	660	12 US-10-072-012-682	Sequence 682, App
18	2930	85.7	672	12 US-10-296-115-1259	Sequence 1259, App
19	2721.5	79.6	677	12 US-10-072-012-683	Sequence 683, App
20	2500	73.1	517	9 US-09-964-277-21	Sequence 21, Appl
21	1326	38.8	625	12 US-10-072-012-699	Sequence 699, App
22	1302	38.1	663	12 US-10-072-012-700	Sequence 700, App
23	1297	37.9	253	15 US-10-108-260A-4872	Sequence 4872, App
24	1075.5	31.5	616	12 US-10-072-012-266	Sequence 266, App
25	917	26.8	501	12 US-10-072-012-702	Sequence 702, App
26	807	23.6	155	9 US-09-964-277-7	Sequence 7, Appl1
27	732.5	21.4	461	12 US-10-072-012-701	Sequence 701, App
28	654.5	19.1	169	14 US-10-346-356-15	Sequence 15, Appl
29	654.5	19.1	170	9 US-09-775-925-26	Sequence 26, Appl1
30	654.5	19.1	170	9 US-09-847-519A-11	Sequence 11, Appl
31	654.5	19.1	170	12 US-10-655-073-16	Sequence 16, Appl
32	654.5	19.1	170	14 US-10-314-058-14	Sequence 14, Appl
33	654.5	19.1	170	14 US-10-405-808-16	Sequence 16, Appl
34	619	18.1	155	9 US-09-964-277-6	Sequence 6, Appl1
35	619	18.1	155	9 US-09-955-732-6	Sequence 6, Appl1
36	469	13.7	444	9 US-09-964-899-47	Sequence 47, Appl
37	469	13.7	482	12 US-10-058-270A-130	Sequence 130, App
38	469	13.7	482	14 US-10-346-356-2	Sequence 2, Appl1
39	469	13.7	482	16 US-10-648-593-164	Sequence 164, App
40	466.5	13.6	381	14 US-10-184-832-2	Sequence 2, Appl1
41	453	13.3	394	9 US-09-736-457-805	Sequence 805, App
42	453	13.3	394	9 US-09-902-941-805	Sequence 805, App
43	453	13.3	394	9 US-09-849-626-805	Sequence 805, App
44	453	13.3	394	12 US-10-283-017-805	Sequence 805, App
45	453	13.3	394	14 US-10-017-754-805	Sequence 805, App

ALIGNMENTS

RESULT 1	
US-09-816-494-2	
; Sequence 2, Application US/09816494	
; Patent No. US20020034807A1	
; GENERAL INFORMATION:	
; APPLICANT: Meyers, Rachel A.	
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY	
; FILE REFERENCE: 10448-030002	
; CURRENT APPLICATION NUMBER: US/09/816,494	
; PRIOR FILING DATE: 2001-03-23	
; PRIOR APPLICATION NUMBER: US 60/191,858	
; NUMBER OF SEQ ID NOS: 10	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 2	
; LENGTH: 665	
; TYPE: PRT	
; ORGANISM: Homo sapiens	
US-09-816-494-2	
QY	Query Match 99.6%; Score 3406; DB 9; Length 665;
	Best Local Similarity 99.7%; Pred. No. 4.6e-250;
	Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY	1 MAHEMIGTQVTRVALLESGTEKVLIDSRFVENVNTSHIDEAININCSKMKRLQQ 60
DB	1 MAHEMIGTQVTRVALLESGTEKVLIDSRFVENVNTSHIDEAININCSKMKRLQQ 60
QY	61 DKVLLTELIGSHAKHVIDIDSGKVVYDQSSQDVASLSDDCFYVLGLKLEKSPNSVHL 120
DB	61 DKVLLTELIGSHAKHVIDIDSGKVVYDQSSQDVASLSDDCFYVLGLKLEKSPNSVHL 120
QY	121 LAGGFAFSPFCFGLCEGKSTLPTICISQPCFLVIANIGPRILPNLYLGGQRYVNLKELI 180
DB	121 LAGGFAFSPFCFGLCEGKSTLPTICISQPCFLVIANIGPRILPNLYLGGQRYVNLKELI 180
QY	181 QNGIGVLAASVTCRPRDPIPSHFLRVVNSFCEKILPMIDKSVDFIEKAKASNGCV 240

```
Db 181 QONGIGYVLAASNTCPKEDFIPESHFLRPVNDSEFCERILPMDKSDVFIKAKASNGCV 240
Qy 241 LVHCLAGISRSATTAIAIYIMKRMMSLDEAVRFYKERRPTISPNFNLGOLLDYEKIKN 300
Db 241 LVHCLAGISRSATTAIAIYIMKRMMSLDEAVRFYKERRPTISPNFNLGOLLDYEKIKN 300
Qy 301 QTGASGPKSKLKLHLEKPNRPVAVSEGGOKSETPLSPCADSATSEAGORPVHPASV 360
Db 301 QTGASGPKSKLKLHLEKPNRPVAVSEGGOKSETPLSPCADSATSEAGORPVHPASV 360
Qy 361 PSVPSVOPSLIEDSPVQALSGHLSDRLSDSNKLRSPSLDIKSYASASMAASLHGF 420
Db 361 PSVPSVOPSLIEDSPVQALSGHLSDRLSDSNKLRSPSLDIKSYASASMAASLHGF 420
Qy 421 SSSSDALEYKPSFTLLDGTNKLCOFSVQELSEQTPETSPDKBEASIPKLTQARPSDSQ 480
Db 421 SSSSDALEYKPSFTLLDGTNKLCOFSVQELSEQTPETSPDKBEASIPKLTQARPSDSQ 480
Qy 481 SKRLHSVRTSSSGTAQKSLSLPLHRSGVEDNHTSFLFGLSTSOOHLTKSAGLGKGMH 540
Db 481 SKRLHSVRTSSSGTAQKSLSLPLHRSGVEDNHTSFLFGLSTSOOHLTKSAGLGKGMH 540
Qy 541 SDILAPQSTPSLTSSWYFATESSHFYASASAIYGSASYSAYSCSOLPTCGDQVYSVRR 600
Db 541 SDILAPQSTPSLTSSWYFATESSHFYASASAIYGSASYSAYSCSOLPTCGDQVYSVRR 600
Qy 601 QKPSDRADSRMSWHEESPFEKQFKRRSCOMEFESIMSENRSHRELKGYGQSSFGSGME 660
Db 601 QKPSDRADSRMSWHEESPFEKQFKRRSCOMEFESIMSENRSHRELKGYGQSSFGSGME 660
Qy 661 IIEVS 665
Db 661 IIEVS 665

RESULT 2
US-09-964-277-2
; Sequence 2, Application US/09964277
; Patent No. US2002013170A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964.277
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-964-277-2
```

```
Query Match 99.6%; Score 3406; DB 9; Length 665;
Best Local Similarity 99.7%; Pred. No. 4.6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 MAHMIQTQVITERVALLESGETEKVLLIDSRPFVEYNTSHLEAININCSKLMKRLQQ 60
Db 1 MAHMIQTQVITERVALLESGETEKVLLIDSRPFVEYNTSHLEAININCSKLMKRLQQ 60
Qy 61 DKVLTIELIHSAXHKVDIDSCQVYVYDQSSQVVASLSDCFVTLLGKLEKSPNSVHL 120
Db 61 DKVLTIELIHSAXHKVDIDSCQVYVYDQSSQVVASLSDCFVTLLGKLEKSPNSVHL 120
Qy 121 LAGGFABFSRCFPGLCGKSTLVPTCISOPCLPVANIGPRTIIPNLYLGGORVLMKEKI 180
Db 121 LAGGFABFSRCFPGLCGKSTLVPTCISOPCLPVANIGPRTIIPNLYLGGORVLMKEKI 180
Qy 181 QONIGIVLAASYCPEDPPIESHFLRPVNDSEFCERILPMDKSDVFIKAKASNGCV 240
Db 181 QONIGIVLAASYCPEDPPIESHFLRPVNDSEFCERILPMDKSDVFIKAKASNGCV 240
```

```
Db 181 QONGIGYVLAASNTCPKEDFIPESHFLRPVNDSEFCERILPMDKSDVFIKAKASNGCV 240
Qy 241 LVHCLAGISRSATTAIAIYIMKRMMSLDEAVRFYKERRPTISPNFNLGOLLDYEKIKN 300
Db 241 LVHCLAGISRSATTAIAIYIMKRMMSLDEAVRFYKERRPTISPNFNLGOLLDYEKIKN 300
Qy 301 QTGASGPKSKLKLHLEKPNRPVAVSEGGOKSETPLSPCADSATSEAGORPVHPASV 360
Db 301 QTGASGPKSKLKLHLEKPNRPVAVSEGGOKSETPLSPCADSATSEAGORPVHPASV 360
Qy 361 PSVPSVOPSLIEDSPVQALSGHLSDRLSDSNKLRSPSLDIKSYASASMAASLHGF 420
Db 361 PSVPSVOPSLIEDSPVQALSGHLSDRLSDSNKLRSPSLDIKSYASASMAASLHGF 420
Qy 421 SSSSDALEYKPSFTLLDGTNKLCOFSVQELSEQTPETSPDKBEASIPKLTQARPSDSQ 480
Db 421 SSSSDALEYKPSFTLLDGTNKLCOFSVQELSEQTPETSPDKBEASIPKLTQARPSDSQ 480
Qy 481 SKRLHSVRTSSSGTAQKSLSLPLHRSGVEDNHTSFLFGLSTSOOHLTKSAGLGKGMH 540
Db 481 SKRLHSVRTSSSGTAQKSLSLPLHRSGVEDNHTSFLFGLSTSOOHLTKSAGLGKGMH 540
Qy 541 SDILAPQSTPSLTSSWYFATESSHFYASASAIYGSASYSAYSCSOLPTCGDQVYSVRR 600
Db 541 SDILAPQSTPSLTSSWYFATESSHFYASASAIYGSASYSAYSCSOLPTCGDQVYSVRR 600
Qy 601 QKPSDRADSRMSWHEESPFEKQFKRRSCOMEFESIMSENRSHRELKGYGQSSFGSGME 660
Db 601 QKPSDRADSRMSWHEESPFEKQFKRRSCOMEFESIMSENRSHRELKGYGQSSFGSGME 660
Qy 661 IIEVS 665
Db 661 IIEVS 665
```

```
RESULT 3
US-10-072-012-680
; Sequence 680, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tcherenev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patuturajan, Meera
; APPLICANT: Shinkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Karazyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072.012
; PRIOR APPLICATION NUMBER: 60/265.102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265.514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265.517
; PRIOR FILING DATE: 2001-01-31
```

PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 680
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-680

Query Match 99.6%; Score 3406; DB 12; Length 665;
Best Local Similarity 99.7%; Pred. No. 4,6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSSTGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSSTGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
QY 61 DKLLITELLOHSAKHKVDIDCSQKVVYDSSQDVASLSDDCLTYLLGLKLEKSFNSVHL 120
DB 61 DKLLITELLOHSAKHKVDIDCSQKVVYDSSQDVASLSDDCLTYLLGLKLEKSFNSVHL 120
QY 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLPVANIGTRILPNLYLGCQDVLMKEIL 180
DB 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLPVANIGTRILPNLYLGCQDVLMKEIL 180
QY 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLPVANIGTRILPNLYLGCQDVLMKEIL 180
DB 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLPVANIGTRILPNLYLGCQDVLMKEIL 180
QY 181 QONGIGYVLAASNTCPKDFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLAASNTCPKDFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRATIAIAYIMKMDMSLDEAYRFVKEKRTISPNFPLQGLDYEKKIKN 300
DB 241 LVHCLAGISRATIAIAYIMKMDMSLDEAYRFVKEKRTISPNFPLQGLDYEKKIKN 300
QY 301 QTASGPKSKLKLHLEKNEPVPAVSEGGQKSETPSPCCADSATSEAGORPVHPASY 360
DB 301 QTASGPKSKLKLHLEKNEPVPAVSEGGQKSETPSPCCADSATSEAGORPVHPASY 360
QY 361 PSVPVQPSLLBDSPLVQALSGHLASDRLEDSNKLKRSFSLDIKSVSYASMAASLHGF 420
DB 361 PSVPVQPSLLBDSPLVQALSGHLASDRLEDSNKLKRSFSLDIKSVSYASMAASLHGF 420
QY 421 SSSBDALFYKSTTLDGTNKLCOFSPVOELSEQTPETSPDKBEASIPKLLQTPARSDSQ 480
DB 421 SSSBDALFYKSTTLDGTNKLCOFSPVOELSEQTPETSPDKBEASIPKLLQTPARSDSQ 480
QY 481 SKRLHSVRTSSSGTAORSLSLPLHRSQSVEDNYHTSFLGLSTSOOHLTKSAGLGKGMH 540
DB 481 SKRLHSVRTSSSGTAORSLSLPLHRSQSVEDNYHTSFLGLSTSOOHLTKSAGLGKGMH 540
QY 541 SDILAPQTSPTSLTSSWYFATSSHPYASAIYGGASAYSAYSCQLPTCGDQVYSVRRR 600
DB 541 SDILAPQTSPTSLTSSWYFATSSHPYASAIYGGASAYSAYSCQLPTCGDQVYSVRRR 600
QY 601 QKSSDADSRSHWESPEFKQKRRSCOMEFESITSEMRSEBELGKVSQSSFSGSME 660
DB 601 QKSSDADSRSHWESPEFKQKRRSCOMEFESITSEMRSEBELGKVSQSSFSGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

DB 661 IIEVS 665

RESULT 4
US-10-168-506-14
Sequence 14, Application US/10168506
Publication No. US20040053229A1
GENERAL INFORMATION:
APPLICANT: PLOMMAN, GREGORY D.
APPLICANT: MARTINEZ, RICARDO
APPLICANT: WHITE, DAVID
APPLICANT: MANNING, GERARD
APPLICANT: SUDASANAM, SUCHA
APPLICANT: HILL, RON
TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
FILE REFERENCE: 038602/1351
CURRENT APPLICATION NUMBER: US/10/168,506
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: PCT/US00/34736
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-10-168-506-14

Query Match 99.6%; Score 3406; DB 12; Length 665;
Best Local Similarity 99.7%; Pred. No. 4,6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSSTGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSSTGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
QY 61 DKLLITELLOHSAKHKVDIDCSQKVVYDSSQDVASLSDDCLTYLLGLKLEKSFNSVHL 120
DB 61 DKLLITELLOHSAKHKVDIDCSQKVVYDSSQDVASLSDDCLTYLLGLKLEKSFNSVHL 120
QY 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLPVANIGTRILPNLYLGCQDVLMKEIL 180
DB 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLPVANIGTRILPNLYLGCQDVLMKEIL 180
QY 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLPVANIGTRILPNLYLGCQDVLMKEIL 180
DB 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLPVANIGTRILPNLYLGCQDVLMKEIL 180
QY 181 QONGIGYVLAASNTCPKDFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLAASNTCPKDFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRATIAIAYIMKMDMSLDEAYRFVKEKRTISPNFPLQGLDYEKKIKN 300
DB 241 LVHCLAGISRATIAIAYIMKMDMSLDEAYRFVKEKRTISPNFPLQGLDYEKKIKN 300
QY 301 QTASGPKSKLKLHLEKNEPVPAVSEGGQKSETPSPCCADSATSEAGORPVHPASY 360
DB 301 QTASGPKSKLKLHLEKNEPVPAVSEGGQKSETPSPCCADSATSEAGORPVHPASY 360
QY 361 PSVPVQPSLLBDSPLVQALSGHLASDRLEDSNKLKRSFSLDIKSVSYASMAASLHGF 420
DB 361 PSVPVQPSLLBDSPLVQALSGHLASDRLEDSNKLKRSFSLDIKSVSYASMAASLHGF 420
QY 421 SSSBDALFYKSTTLDGTNKLCOFSPVOELSEQTPETSPDKBEASIPKLLQTPARSDSQ 480
DB 421 SSSBDALFYKSTTLDGTNKLCOFSPVOELSEQTPETSPDKBEASIPKLLQTPARSDSQ 480
QY 481 SKRLHSVRTSSSGTAORSLSLPLHRSQSVEDNYHTSFLGLSTSOOHLTKSAGLGKGMH 540
DB 481 SKRLHSVRTSSSGTAORSLSLPLHRSQSVEDNYHTSFLGLSTSOOHLTKSAGLGKGMH 540
QY 541 SDILAPQTSPTSLTSSWYFATSSHPYASAIYGGASAYSAYSCQLPTCGDQVYSVRRR 600
DB 541 SDILAPQTSPTSLTSSWYFATSSHPYASAIYGGASAYSAYSCQLPTCGDQVYSVRRR 600

QY 601 QKPSDRADSRMSWHESSPEKOPKRRSCOMERGESIMNSRREELGKVGSQSSPFGSME 660
Db 601 QKPSDRADSRMSWHESSPEKOPKRRSCOMERGESIMNSRREELGKVGSQSSPFGSME 660
QY 661 IIEVS 665
Db 661 IIEVS 665

RESULT 5

US-10-343-357-7
Sequence 7, Application US/10343357
Publication No. US20040058341A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
APPLICANT: YAO, Monique G.; BURFORD, Neil
APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
APPLICANT: GANDHI, Ameena R.; ARVITU, Chandra S.
APPLICANT: LEE, Ernestine A.; HAFALIA, April J. A.
APPLICANT: LU, Dzung Aina M.; TRIBOULEY, Catherine M.
APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
APPLICANT: YUE, Henry; WARREN, Bridget A.
APPLICANT: NGUYEN, Daniel B.; CHAMLA, Narinder K.
APPLICANT: KEARNEY, Liam
TITLE OF INVENTION: PROTEIN PHOSPHATASES
FILE REFERENCE: PI-0173 PCT
CURRENT APPLICATION NUMBER: US/10/343,357
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,679
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/223,272
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/224,309
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 60/226,728
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/229,254
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: US 60/231,366
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CD1
US-10-343-357-7

Query Match 99.6%; Score 3406; DB 12; Length 665;
Best Local Similarity 99.7%; Pred. No. 4,66-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGQVTELVALLSGETEKVLLIDSRPFVEVNTSHIEALININOSKLMKRRLOQ 60
Db 1 MAHEMIGQVTELVALLSGETEKVLLIDSRPFVEVNTSHIEALININOSKLMKRRLOQ 60
QY 61 DKVLITELIHSAGHKVDICSOQVVVYDOSSDVASLSDDCVLVLGKLEKSFNSVHL 120
Db 61 DKVLITELIHSAGHKVDICSOQVVVYDOSSDVASLSDDCVLVLGKLEKSFNSVHL 120
QY 121 LAGGFARFSCFPLGCEGKSTIVPTCISQPCLPVANIIGSTRILPMLYGGCQDVYKELI 180
Db 121 LAGGFARFSCFPLGCEGKSTIVPTCISQPCLPVANIIGSTRILPMLYGGCQDVYKELI 180
QY 181 QONGIGVILNASTYCKRPPIPSHSLRPVNDVDFEKLPLWLDKSVDFIEKAVASNGCV 240
Db 181 QONGIGVILNASTYCKRPPIPSHSLRPVNDVDFEKLPLWLDKSVDFIEKAVASNGCV 240

QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAFRPFYKRPPTISPMNPLGQLLDYEKKIKN 300
Db 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAFRPFYKRPPTISPMNPLGQLLDYEKKIKN 300
QY 301 QTGASGPKSKLKLHLERKNEPVAVSBGQKSETPLSPPCADSATSEAGRPVHPASV 360
Db 301 QTGASGPKSKLKLHLERKNEPVAVSBGQKSETPLSPPCADSATSEAGRPVHPASV 360
QY 361 PSVPSVQPSLLLEDSPVLQALSGHLISADRLSDNSKLRSPSLIKSVYSASMAALHGP 420
Db 361 PSVPSVQPSLLLEDSPVLQALSGHLISADRLSDNSKLRSPSLIKSVYSASMAALHGP 420
QY 421 SSSEDALEYYRSTTLTGNTKLCQPSVQELSTQTEPSPDKBEASIPKKTQTAPSPDSQ 480
Db 421 SSSEDALEYYRSTTLTGNTKLCQPSVQELSTQTEPSPDKBEASIPKKTQTAPSPDSQ 480
QY 481 SKRLHSVRTSSSGTAQRSLSPLRSGSYEDNYHTSFLGLSTSOQHLTKSAGLKGWH 540
Db 481 SKRLHSVRTSSSGTAQRSLSPLRSGSYEDNYHTSFLGLSTSOQHLTKSAGLKGWH 540
QY 541 SDILAPQTSPTSLTSSWYFATBESSHFSASAIYGSASAYSQSLPTCCQDQVYSVRR 600
Db 541 SDILAPQTSPTSLTSSWYFATBESSHFSASAIYGSASAYSQSLPTCCQDQVYSVRR 600
QY 601 QKPSDRADSRMSWHESSPEKOPKRRSCOMERGESIMNSRREELGKVGSQSSPFGSME 660
Db 601 QKPSDRADSRMSWHESSPEKOPKRRSCOMERGESIMNSRREELGKVGSQSSPFGSME 660
QY 661 IIEVS 665
Db 661 IIEVS 665

RESULT 6

US-10-377-072-26
Sequence 26, Application US/10377072
Publication No. US2004009501A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals Inc.
APPLICANT: Curtis, Roy A. J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Glucksmann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Miyoung
TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
FILE REFERENCE: MP103-0180NMIM
CURRENT APPLICATION NUMBER: US/10/377,072
CURRENT FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/995,860
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215,370
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/723,806
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/187,455
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/843,297
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,801
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: US 09/861,801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205,508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/816,494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/815,419

PRIOR FILING DATE: 2001-03-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 665
; TYPE: PRN
; ORGANISM: Homo Sapiens
US-10-377-072-26

Query Match 99.6%; Score 3406; DB 15; Length 665;
Best Local Similarity 99.7%; Pred. No. 4.6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTELVALLSSTGKFKLLIDSRPVEYNTSHILEAININSKLMKRLQ 60
DB 1 MAHEMIGTQVTELVALLSSTGKFKLLIDSRPVEYNTSHILEAININSKLMKRLQ 60
QY 61 DKVLITELIHSKHKVDIDCSQKVVYDQSDQVSLSDCFLTVLLGKLEKSPNSVHL 120
DB 61 DKVLITELIHSKHKVDIDCSQKVVYDQSDQVSLSDCFLTVLLGKLEKSPNSVHL 120
QY 121 LAGGFAFSCPCFGLCEGKSTLVPTCISQPCLPVANIIPTRILPNIYLGQQRVNLKELI 180
DB 121 LAGGFAFSCPCFGLCEGKSTLVPTCISQPCLPVANIIPTRILPNIYLGQQRVNLKELI 180
QY 181 QONGIGVNLASTYCKPDPFIPESHFLRPVNDSECEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGVNLASTYCKPDPFIPESHFLRPVNDSECEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKPTISPNFNLGQLLDYEKKIKN 300
QY 301 QTASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPLSPPCADSATSEAGORVHPASV 360
DB 301 QTASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPLSPPCADSATSEAGORVHPASV 360
QY 361 PSVPSPVPSLLEDSPLVQALSGHLADRLSDNKLKRSFLDIKISVSASMAASLHGF 420
DB 361 PSVPSPVPSLLEDSPLVQALSGHLADRLSDNKLKRSFLDIKISVSASMAASLHGF 420
QY 421 SSSSEDALEYKPSPTLLDGTNKLCOFSPVOELSEOTPESTPDKEEASIPKLLQTPARPSDQ 480
DB 421 SSSSEDALEYKPSPTLLDGTNKLCOFSPVOELSEOTPESTPDKEEASIPKLLQTPARPSDQ 480
QY 481 SKRLHSVTRSSSGTAQRSLSPHRSQSVEDNHTSFLFGLSTSQOHLTKSAGLKGWH 540
DB 481 SKRLHSVTRSSSGTAQRSLSPHRSQSVEDNHTSFLFGLSTSQOHLTKSAGLKGWH 540
QY 541 SDILAPQTSPTSLTSSWYFATESHFSASAIYGGASYSAGSOLPTCGDQVYSVRR 600
DB 541 SDILAPQTSPTSLTSSWYFATESHFSASAIYGGASYSAGSOLPTCGDQVYSVRR 600
QY 601 QKPSDRADSRMSWHEESPPEKQFKRRSCOMEFGESIMSENRREBELGKVSQSSFGSME 660
DB 601 QKPSDRADSRMSWHEESPPEKQFKRRSCOMEFGESIMSENRREBELGKVSQSSFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 7
US-10-257-026-2
; Sequence 2, Application US/10257026
; Publication No. US20040086859A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10K0MS
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07

NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-257-026-2

Query Match 99.6%; Score 3406; DB 16; Length 665;
Best Local Similarity 99.7%; Pred. No. 4.6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTELVALLSSTGKFKLLIDSRPVEYNTSHILEAININSKLMKRLQ 60
DB 1 MAHEMIGTQVTELVALLSSTGKFKLLIDSRPVEYNTSHILEAININSKLMKRLQ 60
QY 61 DKVLITELIHSKHKVDIDCSQKVVYDQSDQVSLSDCFLTVLLGKLEKSPNSVHL 120
DB 61 DKVLITELIHSKHKVDIDCSQKVVYDQSDQVSLSDCFLTVLLGKLEKSPNSVHL 120
QY 121 LAGGFAFSCPCFGLCEGKSTLVPTCISQPCLPVANIIPTRILPNIYLGQQRVNLKELI 180
DB 121 LAGGFAFSCPCFGLCEGKSTLVPTCISQPCLPVANIIPTRILPNIYLGQQRVNLKELI 180
QY 181 QONGIGVNLASTYCKPDPFIPESHFLRPVNDSECEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGVNLASTYCKPDPFIPESHFLRPVNDSECEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKPTISPNFNLGQLLDYEKKIKN 300
QY 301 QTASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPLSPPCADSATSEAGORVHPASV 360
DB 301 QTASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPLSPPCADSATSEAGORVHPASV 360
QY 361 PSVPSPVPSLLEDSPLVQALSGHLADRLSDNKLKRSFLDIKISVSASMAASLHGF 420
DB 361 PSVPSPVPSLLEDSPLVQALSGHLADRLSDNKLKRSFLDIKISVSASMAASLHGF 420
QY 421 SSSSEDALEYKPSPTLLDGTNKLCOFSPVOELSEOTPESTPDKEEASIPKLLQTPARPSDQ 480
DB 421 SSSSEDALEYKPSPTLLDGTNKLCOFSPVOELSEOTPESTPDKEEASIPKLLQTPARPSDQ 480
QY 481 SKRLHSVTRSSSGTAQRSLSPHRSQSVEDNHTSFLFGLSTSQOHLTKSAGLKGWH 540
DB 481 SKRLHSVTRSSSGTAQRSLSPHRSQSVEDNHTSFLFGLSTSQOHLTKSAGLKGWH 540
QY 541 SDILAPQTSPTSLTSSWYFATESHFSASAIYGGASYSAGSOLPTCGDQVYSVRR 600
DB 541 SDILAPQTSPTSLTSSWYFATESHFSASAIYGGASYSAGSOLPTCGDQVYSVRR 600
QY 601 QKPSDRADSRMSWHEESPPEKQFKRRSCOMEFGESIMSENRREBELGKVSQSSFGSME 660
DB 601 QKPSDRADSRMSWHEESPPEKQFKRRSCOMEFGESIMSENRREBELGKVSQSSFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 8
US-10-648-593-240
; Sequence 240, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26

D.Pet X

;; PRIOR APPLICATION NUMBER: 60/406,385
;; PRIOR FILING DATE: 2002-08-27
;; NUMBER OF SEQ ID NOS: 557
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 240
;; LENGTH: 665
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-648-593-240

Query Match 99.6%; Score 3406; DB 16; Length 665;
Best Local Similarity 99.7%; Pred. No. 4.6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIOMSAKHVDIDCSOKVYVYDSSQDVASLSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIOMSAKHVDIDCSOKVYVYDSSQDVASLSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAPFRCFPGLCCEKSTLVPTCISQPCLPVANIPTRIILPNLYGCGQDVLANKELI 180
DB 121 LAGGFAPFRCFPGLCCEKSTLVPTCISQPCLPVANIPTRIILPNLYGCGQDVLANKELI 180
QY 181 QONGIGVYVNASYTCCKPDPFIPESHFLVYVNDSCFKILPWLKSDVPLEKAKANGCV 240
DB 181 QONGIGVYVNASYTCCKPDPFIPESHFLVYVNDSCFKILPWLKSDVPLEKAKANGCV 240
QY 241 LVHCLAGISRATIAIAYIMKMDMSLDEAYRFYKERRPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRATIAIAYIMKMDMSLDEAYRFYKERRPTISPNFNLGQLLDYEKKIKN 300
QY 301 QTGASGPKSKLKLHLKRNPPVAVSEGGQKSTPLSPPCADSAITSEAAQRPVHPASV 360
DB 301 QTGASGPKSKLKLHLKRNPPVAVSEGGQKSTPLSPPCADSAITSEAAQRPVHPASV 360
QY 361 PSVSVOPSLLEDSPVLOVLSGLHLASDRLEDNKLKRSFSLDIKVSYSASMAASLHGF 420
DB 361 PSVSVOPSLLEDSPVLOVLSGLHLASDRLEDNKLKRSFSLDIKVSYSASMAASLHGF 420
QY 421 SSSSEDALEYKPSFTLLDGTNKLCOFSPVOELSHQETPETSBDKEASIPKQLQTARPSDQ 480
DB 421 SSSSEDALEYKPSFTLLDGTNKLCOFSPVOELSHQETPETSBDKEASIPKQLQTARPSDQ 480
QY 481 SKRLHSVRTSSSGTAQRSLSPILHRSGVYEDNYHTSFLFGLSTSOQHLTKSAGLKGWH 540
DB 481 SKRLHSVRTSSSGTAQRSLSPILHRSGVYEDNYHTSFLFGLSTSOQHLTKSAGLKGWH 540
QY 541 SDLIAPQSTPSTLTSSWYFATSSHPYASAIYGSASAYSVCSCOLPTCGDQYVYVRR 600
DB 541 SDLIAPQSTPSTLTSSWYFATSSHPYASAIYGSASAYSVCSCOLPTCGDQYVYVRR 600
QY 601 QKPSRDADRRSRWHESPFKQFKRRSCOMERGESIMSENRRELBELGVQSOSFSFGSME 660
DB 601 QKPSRDADRRSRWHESPFKQFKRRSCOMERGESIMSENRRELBELGVQSOSFSFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 9

US-10-648-593-247
;; Sequence 247, Application US/10648593
;; Publication No. US20040106132A1
;; GENERAL INFORMATION:
;; APPLICANT: Bristol-Myers Squibb Company
;; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
;; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
;; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
;; FILE REFERENCE: D0273 NP

D. Rat

;; CURRENT APPLICATION NUMBER: US/10/648,593
;; CURRENT FILING DATE: 2003-08-26
;; PRIOR APPLICATION NUMBER: 60/406,385
;; PRIOR FILING DATE: 2002-08-27
;; NUMBER OF SEQ ID NOS: 557
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 247
;; LENGTH: 665
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-648-593-247

Query Match 99.6%; Score 3406; DB 16; Length 665;
Best Local Similarity 99.7%; Pred. No. 4.6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIOMSAKHVDIDCSOKVYVYDSSQDVASLSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIOMSAKHVDIDCSOKVYVYDSSQDVASLSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAPFRCFPGLCCEKSTLVPTCISQPCLPVANIPTRIILPNLYGCGQDVLANKELI 180
DB 121 LAGGFAPFRCFPGLCCEKSTLVPTCISQPCLPVANIPTRIILPNLYGCGQDVLANKELI 180
QY 181 QONGIGVYVNASYTCCKPDPFIPESHFLVYVNDSCFKILPWLKSDVPLEKAKANGCV 240
DB 181 QONGIGVYVNASYTCCKPDPFIPESHFLVYVNDSCFKILPWLKSDVPLEKAKANGCV 240
QY 241 LVHCLAGISRATIAIAYIMKMDMSLDEAYRFYKERRPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRATIAIAYIMKMDMSLDEAYRFYKERRPTISPNFNLGQLLDYEKKIKN 300
QY 301 QTGASGPKSKLKLHLKRNPPVAVSEGGQKSTPLSPPCADSAITSEAAQRPVHPASV 360
DB 301 QTGASGPKSKLKLHLKRNPPVAVSEGGQKSTPLSPPCADSAITSEAAQRPVHPASV 360
QY 361 PSVSVOPSLLEDSPVLOVLSGLHLASDRLEDNKLKRSFSLDIKVSYSASMAASLHGF 420
DB 361 PSVSVOPSLLEDSPVLOVLSGLHLASDRLEDNKLKRSFSLDIKVSYSASMAASLHGF 420
QY 421 SSSSEDALEYKPSFTLLDGTNKLCOFSPVOELSHQETPETSBDKEASIPKQLQTARPSDQ 480
DB 421 SSSSEDALEYKPSFTLLDGTNKLCOFSPVOELSHQETPETSBDKEASIPKQLQTARPSDQ 480
QY 481 SKRLHSVRTSSSGTAQRSLSPILHRSGVYEDNYHTSFLFGLSTSOQHLTKSAGLKGWH 540
DB 481 SKRLHSVRTSSSGTAQRSLSPILHRSGVYEDNYHTSFLFGLSTSOQHLTKSAGLKGWH 540
QY 541 SDLIAPQSTPSTLTSSWYFATSSHPYASAIYGSASAYSVCSCOLPTCGDQYVYVRR 600
DB 541 SDLIAPQSTPSTLTSSWYFATSSHPYASAIYGSASAYSVCSCOLPTCGDQYVYVRR 600
QY 601 QKPSRDADRRSRWHESPFKQFKRRSCOMERGESIMSENRRELBELGVQSOSFSFGSME 660
DB 601 QKPSRDADRRSRWHESPFKQFKRRSCOMERGESIMSENRRELBELGVQSOSFSFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 10

US-10-072-012-679
;; Sequence 679, Application US/10072012
;; Publication No. US2004003493A1
;; GENERAL INFORMATION:
;; APPLICANT: Tcherny, Veilizar
;; APPLICANT: Spytek, Kimberly
;; APPLICANT: Zetunen, Bryan

APPLICANT: Patuturajan, Meera
APPLICANT: Shinkete, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Baha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosee, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 679
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-679

Query Match 99.6%; Score 3406; DB 12; Length 690;
Best Local Similarity 99.7%; Pred. No. 4.9e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTRVLVALLESGETEKLVIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
DB 26 MAHEMIGTQVTRVLVALLESGETEKLVIDSRPFVEYNTSHLEAININCSKLMKRLQ 85
QY 61 DKVLITELLQHSKAKHKVDICQKVVVVYDSSQDVASLSDDCTLVLLGLKLESPNSVHL 120
DB 86 DKVLITELLQHSKAKHKVDICQKVVVVYDSSQDVASLSDDCTLVLLGLKLESPNSVHL 145
QY 121 LAGGFAFSCFPGLCGKSTLVPCTISOPCLPVANIGPRLIPNTLYLGGORDVLANKELI 180
DB 146 LAGGFAFSCFPGLCGKSTLVPCTISOPCLPVANIGPRLIPNTLYLGGORDVLANKELI 205
QY 181 QONGIGVILNASTYCPKDFIPBSHFLRVVNDSCFCEKILPWLIDKSVDFTEKASNGCV 240
DB 206 QONGIGVILNASTYCPKDFIPBSHFLRVVNDSCFCEKILPWLIDKSVDFTEKASNGCV 265

QY 241 LVHCLAGISRSATTAIAYIMKMDMSLDEAYRFEKKEPTISPNFNLQGLDYEEKIKXN 300
DB 266 LVHCLAGISRSATTAIAYIMKMDMSLDEAYRFEKKEPTISPNFNLQGLDYEEKIKXN 325
QY 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGGQSETPLSPPCADSATSEAGQRPVHPASV 360
DB 326 QTGASGPKSKLKLHLEKNEPVPVAVSEGGQSETPLSPPCADSATSEAGQRPVHPASV 385
QY 361 PSVPSVQPSLDESPVLVQALSGHLADRLLEDSNKLKSPSLDIKVSYSASMAASLHGF 420
DB 386 PSVPSVQPSLDESPVLVQALSGHLADRLLEDSNKLKSPSLDIKVSYSASMAASLHGF 445
QY 421 SSEDALLEYKPESTTDGNNKLCQFSPVQBLSEQTPETSPDKKEASIPKLTQPARSDSQ 480
DB 446 SSEDALLEYKPESTTDGNNKLCQFSPVQBLSEQTPETSPDKKEASIPKLTQPARSDSQ 505
QY 481 SKRLHSVRTSSSGCTAQRSLSPILHRSQVYEDNHTSFLGLSTSQHLLTKSAGLGKWH 540
DB 506 SKRLHSVRTSSSGCTAQRSLSPILHRSQVYEDNHTSFLGLSTSQHLLTKSAGLGKWH 565
QY 541 SDILAQSTPSTLTSWYFATESHFTYSASAIYGSASYSAYSCQLPTCGDQVSVRR 600
DB 566 SDILAQSTPSTLTSWYFATESHFTYSASAIYGSASYSAYSCQLPTCGDQVSVRR 625
QY 601 QKPSRADRRSRWHESSPEKQPKRRSCOMEFGESIMENSRREELGKYGSSPFGSME 660
DB 626 QKPSRADRRSRWHESSPEKQPKRRSCOMEFGESIMENSRREELGKYGSSPFGSME 685
QY 661 IIEVS 665
DB 686 IIEVS 690

RESULT 11
US-10-072-012-703
Sequence 703, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zethusen, Bryan
APPLICANT: Patuturajan, Meera
APPLICANT: Shinkete, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Baha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosee, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,412
PRIOR FILING DATE: 2001-01-31

;; PRIOR APPLICATION NUMBER: 60/265,395
;; PRIOR FILING DATE: 2001-01-31
;; PRIOR APPLICATION NUMBER: 60/266,406
;; PRIOR FILING DATE: 2001-02-02
;; PRIOR APPLICATION NUMBER: 60/266,767
;; PRIOR FILING DATE: 2001-02-05
;; PRIOR APPLICATION NUMBER: 60/267,057
;; PRIOR FILING DATE: 2001-02-07
;; PRIOR APPLICATION NUMBER: 60/266,975
;; PRIOR FILING DATE: 2001-02-07
;; PRIOR APPLICATION NUMBER: 60/267,459
;; PRIOR FILING DATE: 2001-02-08
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1391
;; SOFTWARE: Patent Ver. 2.1
;; SEQ ID NO: 703
;; LENGTH: 690
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-072-012-703

Query Match 99.6%; Score 3406; DB 12; Length 690;
Best Local Similarity 99.7%; Pred. No. 4.9e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAHEMIGTQIVTERLVALLSEGTKEVLLIDSRPVEYNTSHILEAININCSKIMKRLQ 60
DB 26 MAHEMIGTQIVTERLVALLSEGTKEVLLIDSRPVEYNTSHILEAININCSKIMKRLQ 85
QY 61 DRYLITELIHSARKKVIDCSQKVVYDOSQDVASLSDCFYVLGLKLEKSFNSVHL 120
DB 86 DRYLITELIHSARKKVIDCSQKVVYDOSQDVASLSDCFYVLGLKLEKSFNSVHL 145
QY 121 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGPRLIPNLYLGCQDVANKELM 180
DB 146 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGPRLIPNLYLGCQDVANKELM 205
QY 181 QONGIGVLANSYCCPDPFIPESHFLRVVNDSECEKLLPWLDSVDFIEKAKASNCV 240
DB 206 QONGIGVLANSYCCPDPFIPESHFLRVVNDSECEKLLPWLDSVDFIEKAKASNCV 265
QY 241 LVHCLAGISRSATIALAYIMKRMDSLDEAYRFVVEKRTTISPNNFLGQLLDYEKKIKN 300
DB 266 LVHCLAGISRSATIALAYIMKRMDSLDEAYRFVVEKRTTISPNNFLGQLLDYEKKIKN 325
QY 301 QTGASGPKSKLKLHLKEKNEPVPVAVSEGGQKSETPPLSPCADSATSEAGQRPVHPASV 360
DB 326 QTGASGPKSKLKLHLKEKNEPVPVAVSEGGQKSETPPLSPCADSATSEAGQRPVHPASV 385
QY 361 PSVPVQPSLLEDSPLVQALSGHLHSADRLSDNKLKRSFLDIKISYSASMAASLHGF 420
DB 386 PSVPVQPSLLEDSPLVQALSGHLHSADRLSDNKLKRSFLDIKISYSASMAASLHGF 445
QY 421 SSSSEDALEYKXSTTLDGTNKLCOFSPVOELSEQTPETSPDKEBASIPKQLQTARPSDQ 480
DB 446 SSSSEDALEYKXSTTLDGTNKLCOFSPVOELSEQTPETSPDKEBASIPKQLQTARPSDQ 505
QY 481 SKRLHSVTSSTSSGTAQORSLSPLRSGSVEDNHTSTFLGLSTSQOHLTKSAGLGLKGMH 540
DB 506 SKRLHSVTSSTSSGTAQORSLSPLRSGSVEDNHTSTFLGLSTSQOHLTKSAGLGLKGMH 565
QY 541 SDIILAPQTSPLSTSSWYFATESHFPYASAIYGASASAYSCQLPTCGQOYVSARR 600
DB 566 SDIILAPQTSPLSTSSWYFATESHFPYASAIYGASASAYSCQLPTCGQOYVSARR 625
QY 601 QKPSDRADSRSRWHEBSPFEKQPKRRSCOMEFGEISIMSENRSREBELGKYGSSSFGSME 660
DB 626 QKPSDRADSRSRWHEBSPFEKQPKRRSCOMEFGEISIMSENRSREBELGKYGSSSFGSME 685
QY 661 IIEVS 665
DB 686 IIEVS 690

RESULT 12
US-10-425-114-54204
;; Sequence 54204, Application US/10425114
;; Publication No. US20040034888A1
;; GENERAL INFORMATION:
;; APPLICANT: Liu, Jindong
;; APPLICANT: Zhou, Yihua
;; APPLICANT: Kovalic, David K.
;; APPLICANT: Screen, Steven E
;; APPLICANT: Tabaska, Jack E
;; APPLICANT: Cao, Yongwei
;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;; FILE REFERENCE: 38-21(5313)B
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 73128
;; SEQ ID NO: 54204
;; LENGTH: 690
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
US-10-425-114-54204

Query Match 99.6%; Score 3406; DB 12; Length 690;
Best Local Similarity 99.7%; Pred. No. 4.9e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAHEMIGTQIVTERLVALLSEGTKEVLLIDSRPVEYNTSHILEAININCSKIMKRLQ 60
DB 26 MAHEMIGTQIVTERLVALLSEGTKEVLLIDSRPVEYNTSHILEAININCSKIMKRLQ 85
QY 61 DRYLITELIHSARKKVIDCSQKVVYDOSQDVASLSDCFYVLGLKLEKSFNSVHL 120
DB 86 DRYLITELIHSARKKVIDCSQKVVYDOSQDVASLSDCFYVLGLKLEKSFNSVHL 145
QY 121 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGPRLIPNLYLGCQDVANKELM 180
DB 146 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGPRLIPNLYLGCQDVANKELM 205
QY 181 QONGIGVLANSYCCPDPFIPESHFLRVVNDSECEKLLPWLDSVDFIEKAKASNCV 240
DB 206 QONGIGVLANSYCCPDPFIPESHFLRVVNDSECEKLLPWLDSVDFIEKAKASNCV 265
QY 241 LVHCLAGISRSATIALAYIMKRMDSLDEAYRFVVEKRTTISPNNFLGQLLDYEKKIKN 300
DB 266 LVHCLAGISRSATIALAYIMKRMDSLDEAYRFVVEKRTTISPNNFLGQLLDYEKKIKN 325
QY 301 QTGASGPKSKLKLHLKEKNEPVPVAVSEGGQKSETPPLSPCADSATSEAGQRPVHPASV 360
DB 326 QTGASGPKSKLKLHLKEKNEPVPVAVSEGGQKSETPPLSPCADSATSEAGQRPVHPASV 385
QY 361 PSVPVQPSLLEDSPLVQALSGHLHSADRLSDNKLKRSFLDIKISYSASMAASLHGF 420
DB 386 PSVPVQPSLLEDSPLVQALSGHLHSADRLSDNKLKRSFLDIKISYSASMAASLHGF 445
QY 421 SSSSEDALEYKXSTTLDGTNKLCOFSPVOELSEQTPETSPDKEBASIPKQLQTARPSDQ 480
DB 446 SSSSEDALEYKXSTTLDGTNKLCOFSPVOELSEQTPETSPDKEBASIPKQLQTARPSDQ 505
QY 481 SKRLHSVTSSTSSGTAQORSLSPLRSGSVEDNHTSTFLGLSTSQOHLTKSAGLGLKGMH 540
DB 506 SKRLHSVTSSTSSGTAQORSLSPLRSGSVEDNHTSTFLGLSTSQOHLTKSAGLGLKGMH 565
QY 541 SDIILAPQTSPLSTSSWYFATESHFPYASAIYGASASAYSCQLPTCGQOYVSARR 600
DB 566 SDIILAPQTSPLSTSSWYFATESHFPYASAIYGASASAYSCQLPTCGQOYVSARR 625
QY 601 QKPSDRADSRSRWHEBSPFEKQPKRRSCOMEFGEISIMSENRSREBELGKYGSSSFGSME 660
DB 626 QKPSDRADSRSRWHEBSPFEKQPKRRSCOMEFGEISIMSENRSREBELGKYGSSSFGSME 685

Qy 661 IIEVS 665
Db 686 IIEVS 690

RESULT 13

US-10-072-012-681

; Sequence 681, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernov, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 681
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-072-012-681

Query Match 99.4%; Score 3399; DB 12; Length 665;
Best Local Similarity 99.5%; Pred. No. 1.6e-249;
Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAHEMIGTQVTRVLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
|||||

Db 1 MAHEMIGTQVTRVLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
Qy 61 DKVLITELIOHAKHKVDIDCGOKVYVYDQSSQDVASLSDDCLVTLLKLEKSFNSYHL 120
Db 61 DKVLITELIOHAKHKVDIDCGOKVYVYDQSSQDVASLSDDCLVTLLKLEKSFNSYHL 120
Qy 121 LAGFAEFRCFPGLCCEGKSTLVPTCISQPCLPVANIIGFRLIPNLVLCQDQVNLKELI 180
Db 121 LAGFAEFRCFPGLCCEGKSTLVPTCISQPCLPVANIIGFRLIPNLVLCQDQVNLKELI 180
Qy 181 QONGIGYVNASYTCXKPDFIPESHFLRVVNDVDFCEKILPWLDSKVDPIEKAKANGCV 240
Db 181 QONGIGYVNASYTCXKPDFIPESHFLRVVNDVDFCEKILPWLDSKVDPIEKAKANGCV 240
Qy 241 LVHCLAGISRSATIAAYIMKMDNSLDAAYRVVEXKRTISPENFPLQQLDYEKIKIN 300
Db 241 LVHCLAGISRSATIAAYIMKMDNSLDAAYRVVEXKRTISPENFPLQQLDYEKIKIN 300
Qy 301 QTGASGPKSKLKLHLKENEVPVAVSEGGOKSETPSPCADSATSEAAQRPVPAV 360
Db 301 QTGASGPKSKLKLHLKENEVPVAVSEGGOKSETPSPCADSATSEAAQRPVPAV 360
Qy 361 PSVPSVQPSLLEDSPLVQALSGHLASADRLBDSNKLKRSFSLDIKSVYSASMAASLHG 420
Db 361 PSVPSVQPSLLEDSPLVQALSGHLASADRLBDSNKLKRSFSLDIKSVYSASMAASLHG 420
Qy 421 SSEDALLEYKPSSTLDGNNKLCQSPVQELBQTPETSPDKEASI PKKLQPARSDSQ 480
Db 421 SSEDALLEYKPSSTLDGNNKLCQSPVQELBQTPETSPDKEASI PKKLQPARSDSQ 480
Qy 481 SKRLHSVRTSSGTAQORSLSLPLHRSQVEDVNTSFLGLSTQOHLTKSAGLGLKGMH 540
Db 481 SKRLHSVRTSSGTAQORSLSLPLHRSQVEDVNTSFLGLSTQOHLTKSAGLGLKGMH 540
Qy 541 SDILAQSTPSLTSWYFATSSSHFYASAIYGSASAYSQSLPTCGQDVYVRRR 600
Db 541 SDILAQSTPSLTSWYFATSSSHFYASAIYGSASAYSQSLPTCGQDVYVRRR 600
Qy 601 QKPSDADSRSRHSHSPPEKQPKRRSCOMEGESTMSNRRREELGXGSGSFGSGME 660
Db 601 QKPSDADSRSRHSHSPPEKQPKRRSCOMEGESTMSNRRREELGXGSGSFGSGME 660
Qy 661 IIEVS 665
Db 661 IIEVS 665

RESULT 14

US-10-094-749-2312

; Sequence 2312, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOHYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094,749

```

; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2312
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2312

```

```

Query Match      99.4%; Score 3399; DB 15; Length 665;
Best Local Similarity 99.5%; Pred. No. 1.6e-249;
Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 MAHEMIGTQIVTERLVALLESSTGTEKVLIDSRPVEYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQIVTERLVALLESSTGTEKVLIDSRPVEYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIIOHSAKHVIDIDCSQKVVVYDSSQDVASLSSDCFLVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIIOHSAKHVIDIDCSQKVVVYDSSQDVASLSSDCFLVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCGKSTIVPTCISOPCLPVANIGPRLIPNLVYAGORDVANKELI 180
DB 121 LAGGFAEFSRCFPGLCGKSTIVPTCISOPCLPVANIGPRLIPNLVYAGORDVANKELI 180
QY 181 OONGIGVYLNASTYCPKDPFIPESHFLRVVNDSEFCCKILPWLDSKVDPIEKAKANGCV 240
DB 181 OONGIGVYLNASTYCPKDPFIPESHFLRVVNDSEFCCKILPWLDSKVDPIEKAKANGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRMDSIDEAIRPVKEKRPITISNPNFLGQLLDYEKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKRMDSIDEAIRPVKEKRPITISNPNFLGQLLDYEKIKN 300
QY 301 QTGASGPKSKLHLKEKNEPVPAVSEGGQKSETPLSPCADSATEAAGORPVHPASV 360
DB 301 QTGASGPKSKLHLKEKNEPVPAVSEGGQKSETPLSPCADSATEAAGORPVHPASV 360
QY 361 PSVPSVOPSLIEDSPVQALSGHLNSADRLSDNSKLSFSLDIDSVSYASMAASLHGF 420
DB 361 PSVPSVOPSLIEDSPVQALSGHLNSADRLSDNSKLSFSLDIDSVSYASMAASLHGF 420
QY 421 SSSSDALEYKPSSTTLDGNTKLCQESPVOELSEQTPETSPDKEEASIPKLIQTARPSDSQ 480
DB 421 SSSSDALEYKPSSTTLDGNTKLCQESPVOELSEQTPETSPDKEEASIPKLIQTARPSDSQ 480
QY 481 SKRLHSVTRSSSGTAQRSLSLPLHRSQVEDNYHTSFLGLSTSQOHLIYSAGIGLKGWH 540
DB 481 SKRLHSVTRSSSGTAQRSLSLPLHRSQVEDNYHTSFLGLSTSQOHLIYSAGIGLKGWH 540
QY 541 SDIAPQTSPLTSLTSMWATATSSHPYASAIYGSASYSAYSCSQPTCGDQVYSVRRR 600
DB 541 SDIAPQTSPLTSLTSMWATATSSHPYASAIYGSASYSAYSCSQPTCGDQVYSVRRR 600
QY 601 QRPSPRADSRSRMHEESPFEKQFKRSCCOMFEGSINSEKRSBEIKGKYGSSGSGSME 660
DB 601 QRPSPRADSRSRMHEESPFEKQFKRSCCOMFEGSINSEKRSBEIKGKYGSSGSGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

```

```

RESULT 15
US-10-072-012-258
; Sequence 258, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchearev, Velizar
; APPLICANT: Splet, Kimberly

```

```

; APPLICANT: Zehusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esba
; APPLICANT: Padugaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Futrak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 258
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-258

```

```

Query Match      98.9%; Score 3379.5; DB 12; Length 662;
Best Local Similarity 99.2%; Pred. No. 4.7e-248;
Matches 660; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

```

```

QY 1 MAHEMIGTQIVTERLVALLESSTGTEKVLIDSRPVEYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQIVTERLVALLESSTGTEKVLIDSRPVEYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIIOHSAKHVIDIDCSQKVVVYDSSQDVASLSSDCFLVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIIOHSAKHVIDIDCSQKVVVYDSSQDVASLSSDCFLVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCGKSTIVPTCISOPCLPVANIGPRLIPNLVYAGORDVANKELI 180
DB 121 LAGGFAEFSRCFPGLCGKSTIVPTCISOPCLPVANIGPRLIPNLVYAGORDVANKELI 180
QY 181 OONGIGVYLNASTYCPKDPFIPESHFLRVVNDSEFCCKILPWLDSKVDPIEKAKANGCV 240
DB 181 OONGIGVYLNASTYCPKDPFIPESHFLRVVNDSEFCCKILPWLDSKVDPIEKAKANGCV 240

```

```
Qy 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGOLLJYEKKIXN 300
    |||||
Db 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGOLLJYEKKIXN 300
    |||||
Qy 301 QTGASGPKSLKLIHLEKNEPVPAVSEGGQKSETPLSPCADSATSSEAQRPVHPASV 360
    |||||
Db 301 QTGASGPKSLKLIHLEKNEPVPAVSEGGQKSETPLSPCADSATSSEAQRPVHPA 358
    |||||
Qy 361 PSVPSVQPSLLEDSPVQALSGHLGSADRLSDSNKLRSPFLDIKSVYSASMAASLHGF 420
    |||||
Db 359 -SVPSVQPSLLEDSPVQALSGHLGSADRLSDSNKLRSPFLDIKSVYSASMAASLHGF 417
    |||||
Qy 421 SSSSEDALEYKPBSTTLDGTWKLCQSPVQELSEOTPETSPDKKEASIPKLTOTARPDSQ 480
    |||||
Db 418 SSSSEDALEYKPBSTTLDGTWKLCQSPVQELSEOTPETSPDKKEASIPKLTOTARPDSQ 477
    |||||
Qy 481 SKRLHSVRTSSSGTAQRSLSPHLRSGSVEDNYHTSFLFGLSTSQCHLTRKAGLGLKGMH 540
    |||||
Db 478 SKRLHSVRTSSSGTAQRSLSPHLRSGSVEDNYHTSFLFGLSTSQCHLTRKAGLGLKGMH 537
    |||||
Qy 541 SDILAPQTSPTSLTSSWYFATESSHFYASAGAIYGGASAYSACSQLPFCGDQVYSVRRR 600
    |||||
Db 538 SDILAPQTSPTSLTSSWYFATESSHFYASAGAIYGGASAYSACSQLPFCGDQVYSVRRR 597
    |||||
Qy 601 QKPSDRADSRSRSHESPFEEKQFKRRSCOMEFGESIMSENRSREBLKVGSSQSSFGSGME 660
    |||||
Db 598 QKPSDRADSRSRSHESPFEEKQFKRRSCOMEFGESIMSENRSREBLKVGSSQSSFGSGME 657
    |||||
Qy 661 IIEVS 665
    |||||
Db 658 IIEVS 662
```

Search completed: June 21, 2004, 13:24:57
Job time : 58.0786 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:20:38 ; Search time 22.6939 Seconds

(without alignments)
1512.797 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 3418
Sequence: 1 MAHEMIGTOIVTERLVALLE.....LKVSGSSSRSSSMELIIVS 665Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3406	99.6	665	US-09-816-494-2	Sequence 2, Appl1
2	654.5	19.1	170	US-09-544-716-14	Sequence 14, Appl1
3	654.5	19.1	170	US-09-557-921-15	Sequence 15, Appl1
4	654.5	19.1	170	US-09-564-357-17	Sequence 17, Appl1
5	654.5	19.1	170	US-09-619-380-16	Sequence 16, Appl1
6	469	13.7	482	US-09-557-921-2	Sequence 2, Appl1
7	453	13.3	394	US-09-702-705-805	Sequence 805, App
8	453	13.3	394	US-09-736-457-805	Sequence 805, App
9	453	13.3	394	US-09-614-1248-805	Sequence 805, App
10	453	13.3	394	US-09-671-325-805	Sequence 805, App
11	453	13.3	394	US-09-589-184-805	Sequence 805, App
12	445	13.0	394	US-08-530-290-23	Sequence 23, Appl1
13	445	13.0	394	US-09-702-705-827	Sequence 827, App
14	445	13.0	394	US-09-736-457-827	Sequence 827, App
15	445	13.0	394	US-09-614-1248-827	Sequence 827, App
16	445	13.0	394	US-09-671-325-827	Sequence 827, App
17	445	13.0	394	US-09-589-184-827	Sequence 827, App
18	439.5	12.9	395	US-08-990-379-5	Sequence 5, Appl1
19	433	12.7	367	US-08-990-379-6	Sequence 6, Appl1
20	425	12.4	314	US-09-371-6718-11	Sequence 11, Appl1
21	420	12.3	367	US-08-530-290-24	Sequence 24, Appl1
22	409.5	12.0	313	US-08-990-379-7	Sequence 7, Appl1
23	409	12.0	314	US-09-164-193-22	Sequence 22, Appl1
24	409	12.0	314	US-09-221-448A-22	Sequence 22, Appl1
25	393.5	11.5	393	US-08-990-379-4	Sequence 4, Appl1
26	377	11.0	302	US-09-702-705-806	Sequence 806, App
27	377	11.0	302	US-09-736-457-806	Sequence 806, App

28	377	11.0	302	US-09-614-1248-806	Sequence 806, App
29	377	11.0	302	US-09-671-325-806	Sequence 806, App
30	377	11.0	302	US-09-589-184-806	Sequence 806, App
31	366	10.7	397	US-08-990-379-8	Sequence 8, Appl1
32	346.5	10.1	168	US-09-544-716-13	Sequence 13, Appl1
33	346.5	10.1	168	US-09-557-921-13	Sequence 13, Appl1
34	346.5	10.1	168	US-09-564-357-16	Sequence 16, Appl1
35	346.5	10.1	168	US-09-619-380-15	Sequence 15, Appl1
36	338.5	9.9	170	US-09-544-716-12	Sequence 12, Appl1
37	338.5	9.9	170	US-09-557-921-12	Sequence 12, Appl1
38	338.5	9.9	170	US-09-564-357-15	Sequence 15, Appl1
39	338.5	9.9	170	US-09-619-380-14	Sequence 14, Appl1
40	319	9.3	169	US-09-544-716-16	Sequence 16, Appl1
41	319	9.3	169	US-09-557-921-17	Sequence 17, Appl1
42	319	9.3	169	US-09-564-357-19	Sequence 19, Appl1
43	319	9.3	169	US-09-619-380-18	Sequence 18, Appl1
44	312	9.1	168	US-09-544-716-15	Sequence 15, Appl1
45	312	9.1	168	US-09-557-921-16	Sequence 16, Appl1

ALIGNMENTS

RESULT 1
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-494-2

Query Match	Score	DB 4;	Length	665;
Best Local Similarity	99.6%	Pred. No. 4,8e-288;		
Matches	663;	Conservative	1;	Mismatches 1;
			Indels	0;
			Gaps	0;
QY	1	MAHEMIGTOIVTERLVALLESGTEKYLIDSRFEVYNTSHILEAININCKLMKRLQ	60	
DB	1	MAHEMIGTOIVTERLVALLESGTEKYLIDSRFEVYNTSHILEAININCKLMKRLQ	60	
QY	61	DKVLITELIQAIAKHVDIDCSQKVVVYDQSSQDVASLSDCFLTVLGLTKERSFNSVHL	120	
DB	61	DKVLITELIQAIAKHVDIDCSQKVVVYDQSSQDVASLSDCFLTVLGLTKERSFNSVHL	120	
QY	121	LAGEPAPFRCPCFGLTEBGSSTLVPTCISQPCLEPVANIGPRLIPNLYLGGQRVNLKELI	180	
DB	121	LAGEPAPFRCPCFGLTEBGSSTLVPTCISQPCLEPVANIGPRLIPNLYLGGQRVNLKELI	180	
QY	121	LAGEPAPFRCPCFGLTEBGSSTLVPTCISQPCLEPVANIGPRLIPNLYLGGQRVNLKELI	180	
DB	121	LAGEPAPFRCPCFGLTEBGSSTLVPTCISQPCLEPVANIGPRLIPNLYLGGQRVNLKELI	180	
QY	181	QONGIGVYVNASYTCCKPDPFIPESHFLRVVNDSPFCCKILPMLDKSVDFTEKKAANGCV	240	
DB	181	QONGIGVYVNASYTCCKPDPFIPESHFLRVVNDSPFCCKILPMLDKSVDFTEKKAANGCV	240	
QY	241	LVMCLAGISASATIALAYIMKMDMSLDEAYRVEKRPPTISGNFNLGQLDYKKIKN	300	
DB	241	LVMCLAGISASATIALAYIMKMDMSLDEAYRVEKRPPTISGNFNLGQLDYKKIKN	300	
QY	301	QTASAPKSKLKLHLEKPEPVPAYSEGQKSETPLSPCCASATSEAGORPVHPASV	360	
DB	301	QTASAPKSKLKLHLEKPEPVPAYSEGQKSETPLSPCCASATSEAGORPVHPASV	360	
QY	361	PSVPSVQPSLLEDSPLVQALSGHLASDRLEDSNKLKRSFLDIKSVSYASMAASLHGF	420	

```
Db 361 PSVPSVQPSLIEDSPVQMLSGHLSDRLBDSNKLKRSFSLDIKSYASASMAASHGPF 420
Qy 421 SSSSEDLLEYKPSSTLLDGTNKLQCPSPVQELSEQPEPSPKERASTPKLQTPARPSDQ 480
Db 421 SSSSEDLLEYKPSSTLLDGTNKLQCPSPVQELSEQPEPSPKERASTPKLQTPARPSDQ 480
Qy 481 SKRLHSVPTSSSGTQORSLSPLRSGSVENHTSTSLFGLSTSQOHLTKSAGLGKGMH 540
Db 481 SKRLHSVPTSSSGTQORSLSPLRSGSVENHTSTSLFGLSTSQOHLTKSAGLGKGMH 540
Qy 541 SDIILAPQSTPSLISSWTFATESSHFYASAIYGGSASYSAGSCQLPTCGDQVYSVRR 600
Db 541 SDIILAPQSTPSLISSWTFATESSHFYASAIYGGSASYSAGSCQLPTCGDQVYSVRR 600
Qy 601 QKPSDRADSRSRWHEESPFEPKOPKRRSCQMEFESIMSNRREELGKVSQSSPSGSM 660
Db 601 QKPSDRADSRSRWHEESPFEPKOPKRRSCQMEFESIMSNRREELGKVSQSSPSGSM 660
Qy 661 IIEVS 665
Db 661 IIEVS 665
```

```
RESULT 2
US-09-544-716-14
; Sequence 14, Application US/09544716
; Patent No. 6492157
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; TITLE OF INVENTION: DSP-9 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.415
; CURRENT APPLICATION NUMBER: US/09/544.716
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 14
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-544-716-14
```

```
Query Match 19.1%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 3.8e-49;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Qy 134 GLCEGK-STLVPTCISQPCLPVANIIGPTRILPMLYLGQORVYLNKELIQONGIGYVLNAS 192
Db 1 GLCEGKPAALLPMSLSQPCLPVPSVGLTRILPHLYLGSQKDVINKDMLTONGISYVLNAS 60
Qy 193 YTCPPKDFIPESHFLRPVNDSPCEKILPWLKSVDFTEKAKASNGCVLVHCLAGISRSA 252
Db 61 NSCPKDFICESEFMKVPINDNYCEKILPWLKSTIEFDKAKLSSCQYIVHCLAGISRSA 120
Qy 253 TTAIAIYMKRMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
Db 121 TTAIAIYMKTMGSSDDAYRFVKDRRPSISPNFNLGQLLEYERTLK 167
```

```
RESULT 3
US-09-557-921-15
; Sequence 15, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT APPLICATION NUMBER: US/09/557.921
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-921-15
```

```
Query Match 19.1%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 3.8e-49;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Qy 134 GLCEGK-STLVPTCISQPCLPVANIIGPTRILPMLYLGQORVYLNKELIQONGIGYVLNAS 192
Db 1 GLCEGKPAALLPMSLSQPCLPVPSVGLTRILPHLYLGSQKDVINKDMLTONGISYVLNAS 60
Qy 193 YTCPPKDFIPESHFLRPVNDSPCEKILPWLKSVDFTEKAKASNGCVLVHCLAGISRSA 252
Db 61 NSCPKDFICESEFMKVPINDNYCEKILPWLKSTIEFDKAKLSSCQYIVHCLAGISRSA 120
Qy 253 TTAIAIYMKRMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
Db 121 TTAIAIYMKTMGSSDDAYRFVKDRRPSISPNFNLGQLLEYERTLK 167
```

```
RESULT 4
US-09-564-357-17
; Sequence 17, Application US/09564357
; Patent No. 6645753
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; TITLE OF INVENTION: DSP-5 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.413
; CURRENT APPLICATION NUMBER: US/09/564.357
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 17
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-564-357-17
```

```
Query Match 19.1%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 3.8e-49;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Qy 134 GLCEGK-STLVPTCISQPCLPVANIIGPTRILPMLYLGQORVYLNKELIQONGIGYVLNAS 192
Db 1 GLCEGKPAALLPMSLSQPCLPVPSVGLTRILPHLYLGSQKDVINKDMLTONGISYVLNAS 60
Qy 193 YTCPPKDFIPESHFLRPVNDSPCEKILPWLKSVDFTEKAKASNGCVLVHCLAGISRSA 252
Db 61 NSCPKDFICESEFMKVPINDNYCEKILPWLKSTIEFDKAKLSSCQYIVHCLAGISRSA 120
Qy 253 TTAIAIYMKRMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
Db 121 TTAIAIYMKTMGSSDDAYRFVKDRRPSISPNFNLGQLLEYERTLK 167
```

```
RESULT 5
US-09-619-380-16
; Sequence 16, Application US/09619380
; Patent No. 6649391
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; TITLE OF INVENTION: DSP-11 DUAL SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.418
; CURRENT APPLICATION NUMBER: US/09/619.380
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
```

LENGTH: 170
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-619-380-16

Query Match 19.1%; Score 654.5; DB 4; Length 170;
 Best Local Similarity 72.5%; Pred. No. 3.8e-49;
 Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

QY 134 GLCEGK-STLVPTCISQPCLPVANIGPTRLPLNLYLGCQDVLNKLQNGIGYVLNAS 192
 DB 1 GLCEGKALLPMSLSQPCLPVPSVGLTRILPHLYLGSQDVLNKLQNGIGYVLNAS 60
 QY 193 YTEPKPPIESHPFLRPVNDSPCEKLLPWLKDSVDTEKAKSNGCVLVHCLAGISRSA 252
 DB 61 NSCPKPPICESRMRVPINDNYCEKLLPWLKDSIEFDRAKLSCCQVIVHCLAGISRSA 120
 QY 253 TIAIAYIMKMDSLDEAYRFVKEKPTISPNFNLQQLDYEEKIK 299
 DB 121 TIAIAYIMKMGSSDARFVADRRPSISPNFNLQQLDYERTLK 167

RESULT 6

US-09-557-921-2
 Sequence 2, Application US/09557921
 Patent No. 6551810
 GENERAL INFORMATION:
 APPLICANT: Luche, Ralf M.
 APPLICANT: Wei, Bo
 TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
 FILE REFERENCE: 200125.416
 CURRENT APPLICATION NUMBER: US/09/557,921
 CURRENT FILING DATE: 2000-04-20
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 482
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-557-921-2

Query Match 13.7%; Score 469; DB 4; Length 482;
 Best Local Similarity 35.9%; Pred. No. 3.1e-33;
 Matches 107; Conservative 65; Mismatches 96; Indels 30; Gaps 7;

QY 27 LLIDSRPEVYNTSHILEAININCS-KLMKRRLQDDKVLITELIO-HSAKHVDDICSOX 84
 DB 173 VIIDCRPFMEYNHSHIGAVHINCADKISRRLQGGKITVLDLSCHEGKDSFRRISKE 232
 QY 85 VVYYDSSQDVASLSDCFITVLIGKLEKSFNSVHLLAGFAFSPRCFGLCEBK----- 139
 DB 233 IIVYDENTNEPSRVPESQPLHIVLESIKRGEKELVLKGGISFQKHENMLCDNSLQOE 292
 QY 140 -----STLVPTCISQPCLP-VANIGPTRLPLNLYLGCQDVLNKLQNGIGY 187
 DB 293 CREVGASASASLPLPPI--PTTPDIENBELTFLPLFLGNEQDADDTMQLNIGY 350
 QY 188 VLNAS-----YTCPPKPPIESHPFLRPVNDSPCEKLLPWLKDSVDTEKAKSNGCVLV 242
 DB 351 VIVVTHPLPLVHYEKGJF---NYKRLPATDSNNQNRQYFEEAFETIEBHQGGKLLI 406
 QY 243 HCLAGISRSATIAIAYIMKMDSLDEAYRFVKEKPTISPNFNLQQLDYEEKIK 300
 DB 407 HCQAGVSRSAIVIAIYIMKMTMTDAVYFVKGRPIISPNLNFQQLLEFEDLNN 464

RESULT 7

US-09-702-705-805
 Sequence 805, Application US/09702705
 Patent No. 6504010
 GENERAL INFORMATION:
 APPLICANT: Wang, Tonglong
 APPLICANT: Bangur, Chaitanya S.

APPLICANT: Lodes, Michael A.
 APPLICANT: Fanger, Gary
 APPLICANT: Vedvick, Tom
 APPLICANT: Carter, Darlick
 APPLICANT: Retter, Marc
 APPLICANT: Mannon, Jane
 APPLICANT: Fan, Liqun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121.478C14
 CURRENT APPLICATION NUMBER: US/09/702,705
 CURRENT FILING DATE: 2000-10-30
 NUMBER OF SEQ ID NOS: 1833
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 805
 LENGTH: 394
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-702-705-805

Query Match 13.3%; Score 453; DB 4; Length 394;
 Best Local Similarity 31.4%; Pred. No. 5.5e-31;
 Matches 120; Conservative 73; Mismatches 143; Indels 46; Gaps 13;

QY 19 LESGTEKVLIDSRPEVYNTSHILEAININCSKLMKRRLQDDKVLITELI--OHSAXKH 76
 DB 39 LPSG-GKCLLDPRPLASAGYILGSVNRCTIYRR-AGKSVSLQILPABEVRAR 96
 QY 77 VDDICSOQVVVYDSSQDVASLSDCFITVLIGKLEKSF--NSVHLLAGFAFSPRCFPG 134
 DB 97 LRSGLVSATVYDERSPRAESLREDSVSLVQALRNARBTICLLKGVERFSSEYPE 156
 QY 135 LCEGKSTL-----VPTCISQPCLPVA-----NIGTRILPLNLYLGCQDVLNKL 179
 DB 157 FCSKTRALAIIPPVPSPATEPDLDCSSCGTFLHQEGVELLPFLYLGSAYHARM 216
 QY 180 IQONGIGYVLNASYTCPPKPPIESHPFLRPVNDSPCEKLLPWLKDSVDTEKAKSNGC 239
 DB 217 LDALGITALLNVSDCN-HFEGHYQKCLPVEDNHRKADISSPMETAEIYIDAVKCRGR 275
 QY 240 VLHCLAGISRSATIAIAYIMKMDSLDEAYRFVKEKPTISPNFNLQQLDYEEKIK 299
 DB 276 VLVHCLAGISRSATICLAYIMKMRVRLBEAFVQGRARITISPNFNLQQLDYEEKIK 335
 QY 300 NQTAGSPYSKLLHLKXPNFVPAVBSGQKSETPSPCCADTSKAGQRPVHPAS 359
 DB 336 ATSCAAEAS-----PSGPL-----GERGKTPATP-----TSQVFSFPV-SVG 373
 QY 360 VPSVPSVQPSLSDPLVQALS 381
 DB 374 VHSAPSSLPYL-HSPITTSPS 393

RESULT 8

US-09-736-457-805
 Sequence 805, Application US/09736457
 Patent No. 6509448
 GENERAL INFORMATION:
 APPLICANT: Wang, Tonglong
 APPLICANT: Bangur, Chaitanya S.
 APPLICANT: Lodes, Michael A.
 APPLICANT: Fanger, Gary
 APPLICANT: Vedvick, Tom
 APPLICANT: Carter, Darlick
 APPLICANT: Retter, Marc
 APPLICANT: Mannon, Jane
 APPLICANT: Fan, Liqun
 APPLICANT: Wang, Aijun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121.478C15
 CURRENT APPLICATION NUMBER: US/09/736,457
 CURRENT FILING DATE: 2000-12-13

```
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-805
```

```
Query Match
Best Local Similarity 13.3%; Score 453; DB 4; Length 394;
Matches 120; Conservative 73; Mismatches 143; Indels 46; Gaps 13;
```

```
QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELI--QHSAGHK 76
DB 39 LPSSG-GKCLLDRCRPFLLASAGYILGSAVNRCTIYRRR-AKGSVLEQILPAEEVRRAR 96
QY 77 VDIDCSQKVVYVDQSSQDVASLSDDCLTVLLGLKLEKSF--NSVHLLAGFAEFSCFPG 134
DB 97 LRSGLYSAVIYVDERSPRAESLRDSTVSLVQALRRNAERTDICLKGYERFSSEYBE 156
QY 135 LCEGKSTL-----VPTCISQP-----CLPVA-----NIGPTRILPNLYGCGORDVANKEL 179
DB 157 FCSKTKALAAIPPPVPSPATEPLDLDSCSGTPRHDOEGPVEILPFLYLGSAYHAARDM 216
QY 180 IQONGIGVYLAASYCPKPDPIPESHFLRPVNDSPCEKILPMLDKSVDFTEKAKASNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQYKCIPEVDNHKADISSWPMEAIEYIDAVKDCGR 275
QY 240 VLVHCLAGISRSATITAIYIMKRMDSIDEAYRFVYKERTPTSPNPNFLOLDYEKKIK 299
DB 276 VLVHCOAGISRSATITCLAYIMKKRVLEAFEFVKQRSTISNPFMQQLQFESSQYL 335
QY 300 NOTGASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPLSPPCADATSEAAQGRPVHPAS 359
DB 336 ATSCAAEAS-----PSGFL-----GERGKTPTATP-----TSQPFVFSFPV-SVG 373
QY 360 VPSVPSVQPSLLEDSPLVQALS 381
DB 374 VHSAPSSLPYL--HSPITTPS 393
```

RESULT 9

```
US-09-614-124B-805
; Sequence 805, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B-805
```

Query Match

```
Best Local Similarity 13.3%; Score 453; DB 4; Length 394;
Matches 120; Conservative 73; Mismatches 143; Indels 46; Gaps 13;
```

```
QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELI--QHSAGHK 76
DB 39 LPSSG-GKCLLDRCRPFLLASAGYILGSAVNRCTIYRRR-AKGSVLEQILPAEEVRRAR 96
QY 77 VDIDCSQKVVYVDQSSQDVASLSDDCLTVLLGLKLEKSF--NSVHLLAGFAEFSCFPG 134
DB 97 LRSGLYSAVIYVDERSPRAESLRDSTVSLVQALRRNAERTDICLKGYERFSSEYBE 156
QY 135 LCEGKSTL-----VPTCISQP-----CLPVA-----NIGPTRILPNLYGCGORDVANKEL 179
DB 157 FCSKTKALAAIPPPVPSPATEPLDLDSCSGTPRHDOEGPVEILPFLYLGSAYHAARDM 216
QY 180 IQONGIGVYLAASYCPKPDPIPESHFLRPVNDSPCEKILPMLDKSVDFTEKAKASNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQYKCIPEVDNHKADISSWPMEAIEYIDAVKDCGR 275
```

```
DB 39 LPSSG-GKCLLDRCRPFLLASAGYILGSAVNRCTIYRRR-AKGSVLEQILPAEEVRRAR 96
QY 77 VDIDCSQKVVYVDQSSQDVASLSDDCLTVLLGLKLEKSF--NSVHLLAGFAEFSCFPG 134
DB 97 LRSGLYSAVIYVDERSPRAESLRDSTVSLVQALRRNAERTDICLKGYERFSSEYBE 156
QY 135 LCEGKSTL-----VPTCISQP-----CLPVA-----NIGPTRILPNLYGCGORDVANKEL 179
DB 157 FCSKTKALAAIPPPVPSPATEPLDLDSCSGTPRHDOEGPVEILPFLYLGSAYHAARDM 216
QY 180 IQONGIGVYLAASYCPKPDPIPESHFLRPVNDSPCEKILPMLDKSVDFTEKAKASNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQYKCIPEVDNHKADISSWPMEAIEYIDAVKDCGR 275
QY 240 VLVHCLAGISRSATITAIYIMKRMDSIDEAYRFVYKERTPTSPNPNFLOLDYEKKIK 299
DB 276 VLVHCOAGISRSATITCLAYIMKKRVLEAFEFVKQRSTISNPFMQQLQFESSQYL 335
QY 300 NOTGASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPLSPPCADATSEAAQGRPVHPAS 359
DB 336 ATSCAAEAS-----PSGFL-----GERGKTPTATP-----TSQPFVFSFPV-SVG 373
QY 360 VPSVPSVQPSLLEDSPLVQALS 381
DB 374 VHSAPSSLPYL--HSPITTPS 393
```

RESULT 10

```
US-09-671-325-805
; Sequence 805, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-671-325-805
```

```
Query Match
Best Local Similarity 13.3%; Score 453; DB 4; Length 394;
Matches 120; Conservative 73; Mismatches 143; Indels 46; Gaps 13;
```

```
QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELI--QHSAGHK 76
DB 39 LPSSG-GKCLLDRCRPFLLASAGYILGSAVNRCTIYRRR-AKGSVLEQILPAEEVRRAR 96
QY 77 VDIDCSQKVVYVDQSSQDVASLSDDCLTVLLGLKLEKSF--NSVHLLAGFAEFSCFPG 134
DB 97 LRSGLYSAVIYVDERSPRAESLRDSTVSLVQALRRNAERTDICLKGYERFSSEYBE 156
QY 135 LCEGKSTL-----VPTCISQP-----CLPVA-----NIGPTRILPNLYGCGORDVANKEL 179
DB 157 FCSKTKALAAIPPPVPSPATEPLDLDSCSGTPRHDOEGPVEILPFLYLGSAYHAARDM 216
QY 180 IQONGIGVYLAASYCPKPDPIPESHFLRPVNDSPCEKILPMLDKSVDFTEKAKASNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQYKCIPEVDNHKADISSWPMEAIEYIDAVKDCGR 275
```


[illegible]

```

RESULT 11
US-09-589-184-805
; Sequence 805: Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Panger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRRT
; ORGANISM: Homo sapiens
; US-09-589-184-805

```

Query Match	13.3%	Score 453	DB 41	Length 394
Best Local Similarity	31.4%	Pred.No. 5.5e-31		
Matches 120	Conservative 73	Mismatches 143	Indels 46	Gaps 13

```

Qy 19 LESTGEKVLIDSRPEVNTSHIIIEAININCSKMKRLOODKYLITEL--OHSXKH 76
Db 39 LPBG-GKCLIDRPFPLAHAGYIIIGSVNVRNTIYRR-AKGSVSEQLIPAEYRAR 96
Qy 77 VPIDCSQKVVYDQSSQDAVSLSDCFLLVLLGLKEKSF--NSVHLAGEFAEFSRCPG 134
Db 97 LRSGLSAVIYVDERSPRAESLRBDSTVSLVYQALRNAERTDILCKGGERSRSEYPE 156
Qy 135 LCEGKSTL-----VPTCISQ-----CLPVA-----NIGPRILPNLYLGCDVYLNKL 179
Db 157 FCSKRTALAAIPPPVPSPATPEPLDDSCSCGPTLHDQGPVEIILPFYLGSAVYAAERDM 216
Qy 180 IQONGIGVYLNASTYCPKDPFIPESHFLRPVPNDSCFKIILPWLDSKVFLEKAKANGC 239
Db 217 LDMALGITALLNVASSDDCP--HFEQHGYOKCTIPEVDNHRKADISSWFMEALETVIDAVKDCGR 275
Qy 240 VLVHCLAGISRSATIIAYINRMDSLDEAVRFYKERPTISPNFNLGQLDYEEKIK 299
Db 276 VLVHCGAGISRSKTICLATLMKKKVRLEAEAFYKQGRSTIISPNFSWQGLLQFESQVL 335
Qy 300 NQTGASGPKSKLKLHLEKPNBPVAVSEGGOKSTPLSPCADSATISEAAGRPVHPAS 359
Db 336 ATSCAAEAAS-----PSCPL-----GERGTPATP-----TSQFVFSPFV--SVG 373
Qy 360 VPSVPSVQPSILEDSPLOYALS 381
Db 374 VHSAPSSLPVL--HSPITTSPS 393

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	13.0%;	Score 445;	DB 30;	Length 394;
Best Local Similarity	31.2%;	Pred. No. 2.7e-30;		
Matches 119;	Conservative 73;	Mismatches 144;	Indels 46;	Gaps 13

[illegible]

```
Db 276 VLVHCQAGISRSATICTAYIMMKRVLEAEFEVQORSIISPNSEFWGQLQFESQVL 335
Qy 300 NOTGASGPKSKLHLLEKRNBPVAVSEGGQSETPPLSPCADATSEAAQRPVHPAS 359
Db 336 ATSCAAEAAS-----PSGPL-----RERKKTPTATP-----TSQFVFSPFV-SVG 373
Qy 360 VPSVPSVOPSLLEDSPVLQALS 381
Db 374 VHSAPSSLPYL--HSPITTSPTS 393
```

RESULT 13

```
US-09-702-705-827
; Sequence 827, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriek
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 827
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-705-827
```

Query Match 13.0%; Score 445; DB 4; Length 394;

Best Local Similarity 31.4%; Pred. No. 2,7e-30;

```
Matches 120; Conservative 75; Mismatches 141; Indels 46; Gaps 14;

Qy 19 LESGTEKVLIDSRPVEVNTSHILEAININCSKLMKRLQODKVLITELI--QHSAGKH 76
Db 39 LPSSG-GKCLLDRCRPLAHSAGYILGSVAVRCNTIVRRR-AKGSVSEQLIPAEVEYRAR 96
Qy 77 VDIQSQKVVVYDQSQDVASISSDCFLTVLLGKLEKSF--NSVHLLAGFAEFSRCPFG 134
Db 97 LRSGLYSAVIYDERSPPRAESLREDSTVSLVQALRRNAERTDICLKGGYERFSSEYPE 156
Qy 135 LCEGKSTL-----VPTCISQP-----C-LPVANI-GPTRIILPNLYLGGQRDVNLKEL 179
Db 157 FCSKTKALAIIPPVPSPATPEPLDGCSSCGPPLHDGGPVEILPPLYLGSAYHAARRDM 216
Qy 180 IQQNGIGVYVMASTYCPKPDFIPESHFLRPVNDSCFKILPWLDKSVDFIEKAKASNGC 239
Db 217 LDALGITALLNVSSDCPN-HFEGHYQKCIPEVDNHKADISSWMEAIIEYIDAVKDCRGR 275
Qy 240 VLVHCLAGISRSATICTAYIMKRMDSLDEAYRFYKEKRPITSPNPNFLGQLLDYEKKIK 299
Db 276 VLVHCGAGISRSATICTAYIMMKRVLEAEFEVQORSIISPNSEFWGQLQFESQVL 335
Qy 300 NOTGASGPKSKLHLLEKRNBPVAVSEGGQSETPPLSPCADATSEAAQRPVHPAS 359
Db 336 ATSCAAEAAS-----PSGPL-----RERKKTPTATP-----TSQFVFSPFV-SVG 373
Qy 360 VPSVPSVOPSLLEDSPVLQALS 381
Db 374 VHSAPSSLPYL--HSPITTSPTS 393
```

RESULT 14

```
US-09-736-457-827
; Sequence 827, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriek
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 827
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-827
```

Query Match 13.0%; Score 445; DB 4; Length 394;

Best Local Similarity 31.4%; Pred. No. 2,7e-30;

```
Matches 120; Conservative 75; Mismatches 141; Indels 46; Gaps 14;

Qy 19 LESGTEKVLIDSRPVEVNTSHILEAININCSKLMKRLQODKVLITELI--QHSAGKH 76
Db 39 LPSSG-GKCLLDRCRPLAHSAGYILGSVAVRCNTIVRRR-AKGSVSEQLIPAEVEYRAR 96
Qy 77 VDIQSQKVVVYDQSQDVASISSDCFLTVLLGKLEKSF--NSVHLLAGFAEFSRCPFG 134
Db 97 LRSGLYSAVIYDERSPPRAESLREDSTVSLVQALRRNAERTDICLKGGYERFSSEYPE 156
Qy 135 LCEGKSTL-----VPTCISQP-----C-LPVANI-GPTRIILPNLYLGGQRDVNLKEL 179
Db 157 FCSKTKALAIIPPVPSPATPEPLDGCSSCGPPLHDGGPVEILPPLYLGSAYHAARRDM 216
Qy 180 IQQNGIGVYVMASTYCPKPDFIPESHFLRPVNDSCFKILPWLDKSVDFIEKAKASNGC 239
Db 217 LDALGITALLNVSSDCPN-HFEGHYQKCIPEVDNHKADISSWMEAIIEYIDAVKDCRGR 275
Qy 240 VLVHCLAGISRSATICTAYIMKRMDSLDEAYRFYKEKRPITSPNPNFLGQLLDYEKKIK 299
Db 276 VLVHCGAGISRSATICTAYIMMKRVLEAEFEVQORSIISPNSEFWGQLQFESQVL 335
Qy 300 NOTGASGPKSKLHLLEKRNBPVAVSEGGQSETPPLSPCADATSEAAQRPVHPAS 359
Db 336 ATSCAAEAAS-----PSGPL-----RERKKTPTATP-----TSQFVFSPFV-SVG 373
Qy 360 VPSVPSVOPSLLEDSPVLQALS 381
Db 374 VHSAPSSLPYL--HSPITTSPTS 393
```

RESULT 15

```
US-09-614-124B-827
; Sequence 827, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriek
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:15:52 ; Search time 10.3061 Seconds

(without alignments)

2818.703 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302

Sequence: 1 MAHEMICTQVTERVALLE.....PNNFPLGQLDYERKIKNQF 302

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.78:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	450	28.8	619	2	T15969
2	425	27.2	367	1	S24411
3	423	27.0	394	2	A56115
4	418	26.7	314	1	A57126
5	415	26.5	367	1	S29090
6	415	26.5	367	2	S52265
7	409	26.2	314	2	B57126
8	377.5	24.1	393	2	A56947
9	351	22.4	384	1	I38890
10	319	20.4	303	2	T46405
11	289	18.5	365	2	T32494
12	223.5	14.3	330	2	T39698
13	222.5	14.2	186	2	T16056
14	212.5	13.6	226	2	T21380
15	208	13.3	272	2	T18915
16	207.5	13.3	220	2	JC7885
17	205	13.1	223	2	I49365
18	202	12.9	364	1	S31304
19	201.5	12.9	142	2	T03074
20	195.5	12.5	185	1	A47126
21	195	12.5	283	2	G84458
22	193	12.3	205	2	I49364
23	185	11.8	807	1	S44538
24	183.5	11.7	276	2	T46906
25	181.5	11.6	278	2	T39517
26	179.5	11.5	866	2	F88481
27	173.5	11.1	204	2	T17802
28	171	10.9	489	1	S58725
29	162	10.4	771	2	T47666

30	160.5	10.3	580	2	T18439	hypothetical prote
31	160.5	10.3	600	2	T18446	hypothetical prote
32	157	10.0	209	1	S48459	probable dual spec
33	152	9.7	292	2	S41012	hypothetical prote
34	150.5	9.6	169	2	T30684	probable dual spec
35	147.5	9.4	597	1	S43743	probable dual spec
36	145.5	9.3	171	1	I36845	dual specificity p
37	145.5	9.3	171	2	T28522	probable dual spec
38	145.5	9.3	171	2	B72161	probable dual spec
39	144.5	9.2	171	1	Q0V2H1	dual specificity p
40	138.5	8.9	171	1	A42514	dual specificity p
41	134	8.6	150	2	T21489	hypothetical prote
42	134	8.6	272	2	T19418	hypothetical prote
43	127.5	8.2	171	1	B47452	dual specificity p
44	121	7.7	928	2	S50578	hypothetical prote
45	117.5	7.5	160	2	T10278	protein tyrosine p

ALIGNMENTS

RESULT 1

T15969

hypothetical protein F08B1.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15969

R:Chisoe, S.

submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid F08B1.

A:Reference number: Z18439

A:Accession: T15969

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-619 <CHI>

A:Cross-references: EMBL:U23178; NID:9726421; PID:9726422; PIDN:AAC46719.1; CESP:F08B1.1

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:F08B1.1

A:Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Query Match 28.8%; Score 450; DB 2; Length 619;
Best Local Similarity 50.6%; Pred. No. 2.3e-29;
Matches 90; Conservative 34; Mismatches 48; Indels 6; Gaps 4;

QY	123	GGPAERSRCPPGLCEGKS--TLVPTCISQPC--PVANIGPTRILPNLYIGCQPDVANK	178
DB	99	GGFKQFAQYQPOLCESESGMTRLPQSLQPCLSQPTGD-GITLTPMIVIGSQIDSDT	157
QY	179	LIQONGIGVYVNASYTPKPKDFIPE-SHFLRVPNVDSFCFKILPMIDKSVDFIKAVSN	237
DB	158	MLDLADISVINISMTPKVCIKEDNFMKIPNDYSQELSYFPMAYEFLEKCRAG	217
QY	238	GCVLVHCLAGISRSATIAIVIMKRMDSIDEAVRFVKEKRPRTSPNFFLGQLDYER	295
DB	218	KKCLIHCLAGISRSPTIAISYIMRYMKGSDDAIVRYKERRPSIPNFFMGQLDYER	275

RESULT 2

S24411

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - mouse

N:Alternate names: 3CH134 protein; protein-tyrosine-phosphatase erp, nonreceptor type 10

C:Species: Mus musculus (house mouse)

C:Date: 19-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999

C:Accession: A54681; S24411

R:Noguchi, T.; Metz, R.; Chen, L.; Mattei, M.G.; Carrasco, D.; Bravo, R.

Mol. Cell. Biol. 13, 5195-5205, 1993

A:Title: Structure, mapping, and expression of erp, a growth factor-inducible gene encod

A:Reference number: A54681; PMID:8355678

A:Accession: A54681

A:Molecule type: DNA

A:Residues: 1-367 <NOG>

A:Cross-references: GB:S64851; NID:g409976; PIDN:AAB27882.1; PID:g409977

R;Charles, C.H.; Abler, A.S.; Lau, L.F.
Oncogene 7, 187-190, 1992

A:Title: cDNA sequence of a growth factor-inducible immediate early gene and characteriz

A:Reference number: S24411; MUID:92158357; PMID:1741163

A:Accession: S24411

A:Molecule type: mRNA

A:Residues: 1-367 <CHAS>

A:Cross-references: EMBL:X61940; NID:949735; PID:949736

C:Genetics:

A:Gene: erp

A:Introns: 123/1; 172/1; 245/1

C:Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity

C:Keywords: immediate-early protein; phosphoprotein; phosphoric monoester hydrolase

F:181-312/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

F:256/Active site: Cys (phosphocysteine intermediate) #status predicted

F:264/Binding site: substrate phosphate (Arg) #status predicted

Query Match

Best Local Similarity 27.2%; Score 425; DB 1; Length 367;
Matches 104; Conservative 48; Mismatches 129; Indels 22; Gaps 5;

```

QY 15 LVALLSGTEKVLIDSRPFVEYNTSHILEAININCKMKRRLOQDKVLTITELIQHSK 74
Db 13 LRLALREGAQCCLLDRCSPFAPNACHIVNVPSTVRRRAKMGLEHIVPAELR 72
QY 75 HKVDIDCSQKVVYVYDSSQDVASLSDCFVLTLGKLT--EKSFNVSVHLAAGFAEFRCF 132
Db 73 GRLLAGVAVVLDLDRSASLDGAKRDGTLALAGALCREARSTQVFLLGGVHARSASC 132
QY 133 PGLCESKSTLVPTGISQP-----CLPVANI-----GPTRLIPNLVYIGCQRDVY 175
Db 133 PELCSQST--PTGLSTPLSTSVDPDSASGSSCSTPLVYDGGVELLSTFLYGSAYHAS 190
QY 176 NKEILIQNGIGVNLASVYTCRPPDFIESHFLRVVPVDSFCEKILPWLKSVDFIKAKA 235
Db 191 RQMDALGITALIVNSANCEN-HFEHGYQKSPVVDNHRKADISSWFMEALDFIDSTID 249
QY 236 SNGCVLVHCLAGISRSATTAIYIMKRMDSLDEAVRFVKEKPTISPNFNLGQLDYE 295
Db 250 AGGRVFNVCQAGISRSATICLAYLMRNVKLDLDEAFVQKRSITISPNFSGQLDQFE 309
QY 296 KKI 298
Db 310 SQV 312

```

RESULT 3

A56115

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 4 - human

N:Alternate names: dual specificity phosphatase HVH2

C:Species: Homo sapiens (man)

C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Apr-1998

C:Accession: A56115

R:Guan, K.L.; Burch, R.

J. Biol. Chem. 270, 7197-7203, 1995

A:Title: Isolation and characterization of a novel dual specific phosphatase, HVH2, whic

A:Reference number: A56115; MUID:95221370; PMID:7555768

A:Accession: A56115

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-394 <GUA>

A:Cross-references: GB:U21108

C:Genetics:

A:Gene: GDB:DUSP4; HVH2; MKP-2

A:Cross-references: GDB:433893

A:Map position: 8p21-8p11.2

C:Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity

C:Keywords: phosphoprotein; phosphoric monoester hydrolase

F:203-334/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

F:280/Active site: Cys (phosphocysteine intermediate) #status predicted

F:286/Binding site: substrate phosphate (Arg) #status predicted

Query Match

27.0%; Score 423; DB 2; Length 394;

Best Local Similarity 33.4%; Pred. No. 2.2e-27;
Matches 100; Conservative 63; Mismatches 114; Indels 22; Gaps 9;

```

QY 19 LBSGTEKVLIDSRPFVEYNTSHILEAININCKMKRRLOQDKVLTITELI--QHSARK 76
Db 39 LPBG--GKCLLDRCPLAHSAGYILGVSVNRNTVRR--AKGSVLEQILPAEEVRR 96
QY 77 VDDIDCSQKVVYVYDSSQDVASLSDCFVLTLGKLEKSF--NSVHLAAGFAEFRCFPG 134
Db 97 LBSGLVSAIVYVDSPPRAESLREDSVSLVQALRRNAERTDICLKGGERRSSRYE 156
QY 135 LBSGKSTL-----VPTGISQP-----C-LPVANI--GPTRLIPNLVYIGCQRDVANKEL 179
Db 157 FCSKTKVLAALVPPVPSPATPEPLDLCSSCGTPIHDGCGVELLPFLYGSAYHAARRDM 216
QY 180 IQONGIGVNLASVYTCRPPDFIESHFLRVVPVDSFCEKILPWLKSVDFIKAKASNGC 239
Db 217 LDALGITLALVSSDCNR-HFEHGYQKCIPEVDNHRKADISSWFMEALIEYDAVDCRR 275
QY 240 VLVHCLAGISRSATTAIYIMKRMDSLDEAVRFVKEKPTISPNFNLGQLDYEKTI 298
Db 276 VLVHQCAGISRSATICLAYLMKRVLRBEAFVQKRSITISPNFSGQLDQFEQV 334

```

RESULT 4

A57126

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 2 - human

N:Alternate names: mitogen-induced nuclear tyrosine phosphatase; phosphatase of activat-

C:Species: Homo sapiens (man)

C>Date: 03-Nov-1995 #sequence_revision 02-Jul-1996 #text_change 11-Jun-1999

C:Accession: A57126

R:Roman, P.J.; Davis, P.; Moskaluk, C.A.; Keane, M.; Krutzsch, H.; Siebenlist, U.; Keil

Science 259, 1763-1766, 1993

A:Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.

A:Reference number: A57126; MUID:93206122; PMID:7681221

A:Accession: A57126

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-314 <ROH>

A:Cross-references: GB:L11329; NID:9559539; PID:AAA50779.1; PID:g292376

C:Genetics:

A:Gene: GDB:DUSP2

A:Cross-references: GDB:139200

A:Map position: 2q11-2q11

C:Function:

A:Description: catalyzes the hydrolysis of peptidyl-phosphoserine, -phosphothreonine, and

C:Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity

C:Keywords: nucleus; phosphoprotein; phosphoric monoester hydrolase

F:140-160/Region: nuclear location signal

F:180-311/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

F:257/Active site: Cys (phosphocysteine intermediate) #status predicted

F:263/Binding site: substrate phosphate (Arg) #status predicted

Query Match

Best Local Similarity 26.7%; Score 418; DB 1; Length 314;
Matches 109; Conservative 45; Mismatches 105; Indels 44; Gaps 8;

```

QY 24 EKVLLIDSRPFVEYNTSHILEAININCKMKRR-----LOQDKVLTITELIQHS 72
Db 25 ERTLLIDCRPLAFGRHVRARVPVNALRRRRRGPAAVLACLLPDRALRTRLVGE 84
QY 73 AKHKVDIDCSQKVVYVYDSSQDVASLSD--CFVLTLGKLEKSFNSVHLAAGFAEF 128
Db 85 L-----AAAVLDEGSASVALRDPSPAHVLAALHETRAGTAVYFRGGDGF 135
QY 129 SBCFGLC-EKGSSTLV-----TCISQCPVPANI--GPTRLIPNLVYIGCQRDVANKEL 181
Db 136 QCCCPDLCSAPALPPTGKTRSDSRAPVVDGGPVILTYLFLGSSSHSDLDQGIQ 195
QY 182 QNGICVNLASVYTCRPPDFIESHF-----LRPVNDSFCEKILPWLKSVDFIEKAKA 235
Db 196 ACGITAVNLVASCP-----NHFEGLFRYKSIPEVDNQMWEISNMFQELGFIQWVKN 248

```

QY 236 SNGCVLVHCLAGISRSATIAIAYIMKMDVSLDEAYRFEVKEKPTISPENFPGQLLDYE 295
 DB 249 SGRRLVHCOAGISRSATICTALIMQSRRLDEAPFVQKRGVISPNSFMGQLLQFE 308
 QY 296 KK1 298
 DB 309 TQV 311

RESULT 5

S29090
 dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human
 N/Alternate names: protein-tyrosine-phosphatase C1100; protein-tyrosine-phosphatase, not
 C/Species: Homo sapiens (man)
 C/Date: 25-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999
 C/Accession: S29090; A53052
 R/Keyes, S.M.; Emalle, E.A.
 Nature 359, 644-647, 1992
 A/Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine
 A/Reference number: S29090; MUID:93024952; PMID:1406996
 A/Accession: S29090
 A/Molecule type: mRNA
 A/Residues: 1-367 <KEY>
 A/Cross-references: EMBL:X68277; NID:G29980; PIDN:CA448338.1; PID:G29981
 R/Kwak, S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.
 J. Biol. Chem. 269, 3596-3604, 1994
 A/Title: Isolation and characterization of a human dual specificity protein-tyrosine pho
 A/Reference number: A53052; MUID:94148864; PMID:8106404
 A/Accession: A53052
 A/Molecule type: DNA
 A/Residues: 1-367 <KWA>
 A/Experimental source: leukocyte
 A/Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804,
 C/Genetics:
 A/Gene: GDB:DUSP1; PTPN10
 A/Cross-references: GDB:136197; OMIM:600714
 A/Map position: 5q34-5q34
 C/Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity
 C/Keywords: heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced
 F/181-312/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>
 F/258/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 26.5%; Score 415; DB 1; Length 367;
 Best Local Similarity 33.9%; Pred. No. 9.2e-27;
 Matches 102; Conservative 47; Mismatches 134; Indels 18; Gaps 4;

QY 15 LVALLBSGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELIQHSK 74
 DB 13 LRLALLGERAAQCILLDCRSFFAFNAGHIAQSVNRSTIVRRRAKMGLEHIVPNAELR 72
 QY 75 HKYVDICQKVVVYDQSSQDVASLSDCFLTVLGLK--EKSFNSVHLAAGFAEFGRCF 132
 DB 73 GRLLAGVHVAHVLLDERSAALDGAKRQDGTALAAAGALCREARAQVFFLGGYEAFASGC 132
 QY 133 PGICBEGK-----STVPTCISQCLPVANI-----GPTRLPMLVLYGCRDVLNK 177
 DB 133 PELCSKQSTPMGLSLPLSTSVPSAEGSCSSCPPLYDQGVETLIPFLVLSAYHNASRK 192
 QY 178 ELIQONGIGVVLNASTYCPKPDPIESHFLRPVNDSCFKILPWLDSVDFTEKAKASN 237
 DB 193 DMDALGITLALINVSANCPN-HFEGHYQKSIPEVDNKHADISWFEALDFIDSIKAG 251
 QY 238 GCYLVHCLAGISRSATIAIAYIMKMDVSLDEAYRFEVKEKPTISPENFPGQLLDYEKK 297
 DB 252 GRVFEVHCQAGISRSATICTALVIMKTRVVKLDEAFEVQKRSIISPNSFMGQLLQFE 311
 QY 298 I 298
 DB 312 V 312

RESULT 6

S52265
 dual specificity phosphatase (EC 3.1.3.-) 1 - rat
 N/Alternate names: protein-tyrosine-phosphatase C1100; protein-tyrosine-phosphatase, noni
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999
 C/Accession: S52265
 R/Muda, M.; Schlegel, W.; Arkinstall, S.
 submitted to the EMBL Data Library, January 1995
 A/Description: Pathways regulating C1100 gene expression in pituitary cells.
 A/Reference number: S52265
 A/Accession: S52265
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-367 <KUD>
 A/Cross-references: EMBL:X64004; NID:G642264; PIDN:CA58828.1; PID:G642265
 C/Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity
 C/Keywords: phosphoprotein; phosphoric monoester hydrolase
 F/181-312/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>
 F/258/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 26.5%; Score 415; DB 2; Length 367;
 Best Local Similarity 34.0%; Pred. No. 9.2e-27;
 Matches 104; Conservative 48; Mismatches 126; Indels 28; Gaps 6;

QY 15 LVALLBSGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELI 69
 DB 13 LRLALLGERAAQCILLDCRSFFAFNAGHIAQSVNRSTIVRRRAKMGLEHIVPNTL- 71
 QY 70 QHSKHKVDICQKVVVYDQSSQDVASLSDCFLTVLGLK--EKSFNSVHLAAGFAE 127
 DB 72 ----RRLLAGVHVAHVLLDERSAALDGAKRQDGTALAAAGALCREARSTGVFFLGGYEA 127
 QY 128 FSCPCFGLCBEGK-----STVPTCISQCLPVANI-----GPTRLPMLVLYGCR 172
 DB 128 FSASCPCLCSKQSTPMGLSLPLSTSVPSAEGSCSSCPPLYDQGVETLIPFLVLSAY 187
 QY 173 DVNLKELIQONGIGVVLNASTYCPKPDPIESHFLRPVNDSCFKILPWLDSVDFTEK 232
 DB 188 HASKRMIDLALGITLALINVSANCPN-HFEGHYQKSIPEVDNKHADISWFEALDFID 246
 QY 233 AKASNCVLVHCLAGISRSATIAIAYIMKMDVSLDEAYRFEVKEKPTISPENFPGQL 292
 DB 247 IKDAGRVFHCQAGISRSATICTALVIMKTRVVKLDEAFEVQKRSIISPNSFMGQL 306
 QY 293 DYKKT 298
 DB 307 QFESQV 312

RESULT 7

B57126
 dual specificity phosphatase (EC 3.1.3.-) 2 - mouse
 N/Alternate names: mitogen-induced nuclear protein-tyrosine-phosphatase PAC-1
 C/Species: Mus musculus (house mouse)
 C/Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 24-Apr-1998
 C/Accession: B57126
 R/Rohan, P.J.; Davis, P.; Moskaluk, C.A.; Kearns, M.; Krutzsch, H.; Siebenlist, U.; Kelly
 Science 259, 1763-1766, 1993
 A/Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
 A/Reference number: A57126; MUID:93206122; PMID:7681221
 A/Accession: B57126
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-314 <ROH>
 A/Cross-references: GB:111330
 C/Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity
 C/Keywords: nucleus; phosphoprotein; phosphoric monoester hydrolase
 F/180-311/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>
 F/257/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/263/Binding site: substrate phosphate (Arg) #status predicted

Query Match 26.2%; Score 409; DB 2; Length 314;

Best Local Similarity 36.0%; Pred. No. 2,4e-26;
Matches 107; Conservative 45; Mismatches 113; Indels 32; Gaps 7;

```
QY 24 EKVLLIDSRPVEYNTSHILEAININCSKLMKR-----LOODKVLITELIQHS 72
Db 25 EKTLLIDCRPFLAFCSHYRAARPPVMAALLRRRAPGTPAALACLIPDALRLRGE 84
QY 73 AKHKVDIDCSQKVYVYDSSQDVASLSDCFLVTLKLEKSF---NSVHLIAGGAEP 128
Db 85 L-----ARAVVLDESSASVTELPDPGPAHLLALQEMKRGPTTCFLRGKFS 135
QY 129 SRCPGLC-EKSKTLVP-----TCISQCLPVPANI-GPTRIIPLYGCGQDVLNKLK 181
Db 136 QTYCPDLCSFAPDAALPFAGAENSNDRPVLYDQSGPVELIPVLYIGSCNHSDDIQGL 195
QY 182 QNGIGVYVNASTYCPKDPFIPESHFLRPVNDSCFKILPWLDKSDVFIKAVANGCYL 241
Db 196 ACGITAVLVNASCPN-HFEGLFHYKSIPEVDNQWVSIAMFQDAISFIDSVKNSGRVL 254
QY 242 VHLIAGISRSATTAIAYIMKMDSLDEAYRFVYKERTSPNPFGLQDLYEKKI 298
Db 255 VHCQAGISRSATICTAVLIQSHRILDEADPFYKRGVISPWFSGQLQLETGV 311
```

RESULT 8

dual specificity phosphatase (EC 3.1.3.-) HVH2 - rat
N/Alternate names: mtogen-activated protein kinase phosphatase 2
C/Species: Rattus norvegicus (Norway rat)
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #ext_change 17-Mar-1999
C/Accession: A56947
R/Miscr-Press, A.; Rim, C.S.; Yao, H.; Robertson, M.S.; Stork, P.J.S.
J. Biol. Chem. 270, 14587-14596, 1995
A/Title: A novel mtogen-activated protein kinase phosphatase. Structure, expression, and
A/Reference number: A56947; MUID:95301550; PMID:7782322
A/Accession: A56947
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-393 <WIS>
C/Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
C/Keywords: phosphoprotein, phosphoric monoester hydrolase
F/202-333/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F/279/Active site: Cys (phosphocysteine intermediate) #status predicted
F/285/Binding site: substrate phosphate (arg) #status predicted

Query Match

Best Local Similarity 24.1%; Score 377.5; DB 2; Length 393;
Matches 95; Conservative 54; Mismatches 124; Indels 33; Gaps 8;

```
QY 16 VALLESGETEVLIDSRPVEYNTSHILEAININCSKLMKRLOODKVLITELIQSAKH 75
Db 38 IGLLSGG--KCLLDRCRFLAHSAQYINGVNVRCNTIVRR-AGGSVLEQIL--PAEE 92
QY 76 KV-----DIDCSQKVYVYDSSQDVASLSDCFLVTLKLEKSFNSVHLIAGGAEP 129
Db 93 EKRPCALASTRLSSSTMTQPARRESPGQHSVAGRALRRNAET--DICLLGGYRRFS 150
QY 130 RCFPLCGKSTVPTICISQCLPVPANI-----GPTRIIPLYGCGQDVLNKLK 172
Db 151 SEYEPFCGKTAAL--AALPPVPSTNESLDLGSSCGTLPDQGGPVELIPVLYIGSAV 208
QY 173 DVLNKLIIQNGIGVYVYDSCPKDPFIPESHFLRPVNDSCFKILPWLDKSDVFIK 232
Db 209 HAARDMDLDALGITLVNVSDDCN-HFEGHYQKCIPEVDNKKADISSFMEALTEYDA 267
QY 233 AKANGCVLVHCLAGISRSATTAIAYIMKMDSLDEAYRFVYKERTSPNPFGLQDLY 292
Db 268 VKDRCGVLVHCGGISRSATICTAVLMKRVLEBAEFVYQRRSISIPNFSFGQL 327
QY 293 DYKKKI 298
Db 328 QFESQV 333
```

RESULT 9

138890

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 5 - human
N/Alternate names: dual specificity phosphatase B23; dual-specificity phosphatase hvh-3,
C/Species: Homo sapiens (man)
C/Date: 16-Feb-1996 #sequence_revision 01-Mar-1996 #ext_change 11-Jun-1999
C/Accession: 138890; A55313
J./Kwak, S.P.; Dixon, J.E.
J. Biol. Chem. 270, 1156-1160, 1995
A/Title: Multiple dual specificity protein tyrosine phosphatases are expressed and regulat
A/Reference number: A55432; MUID:95138103; PMID:7836374
A/Accession: 138890
A/Molecule type: mRNA
A/Residues: 1-384 <RES>
A/Cross-references: EMBL:U6966; NID:9642012; PIDD:AMB06261.1; PID:9642013
A/Experimental source: placenta
R/Ishibashi, T.; Bottaro, D.P.; Michiel, P.; Kelley, C.A.; Aaronson, S.A.
J. Biol. Chem. 269, 29897-29902, 1994
A/Title: A novel dual specificity phosphatase induced by serum stimulation and heat shock
A/Reference number: A55313; MUID:95050849; PMID:7961985
A/Accession: A55313
A/Molecule type: mRNA
A/Residues: 1-8, 'GHV', 12-70, 'R', 72-104, 'F', 107-362, 'RCLPTQSSSAEALMORPNPAKTGMESAOPOQL'
A/Cross-references: GB:U15932; NID:9606971; PIDD:AAA64693.1; PID:9606972
A/Experimental source: mammary epithelial cells
C/Genetics:
A/Gene: GDB:DUSP5
A/Cross-references: GDB:385447
A/Map position: 10q25-10q25
C/Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
C/Keywords: phosphoprotein, phosphoric monoester hydrolase
F/186-317/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F/263/Active site: Cys (phosphocysteine intermediate) #status predicted
F/269/Binding site: substrate phosphate (arg) #status predicted

Query Match

Best Local Similarity 22.4%; Score 351; DB 1; Length 384;
Matches 94; Conservative 54; Mismatches 124; Indels 48; Gaps 7;

```
QY 18 LIESGETEVLIDSRPVEYNTSHILEAININCSKLMKRLOODKVLITELIQSAKHV 77
Db 15 LRKEAARCVVLDGRPYLAFAASNVRSGLNVNVSIVLRARGAV-----SARYVL 66
QY 78 DIDCG-----KVYVYDSSQDVASLSDCFLVTLKLEKSFNS--VHLIA 122
Db 67 PDEAARLILDEGGGVAVVLLDQSRHMQKREESAARVLTSLACLIPAGRYVFLK 126
QY 123 GGAFAFSRCFPLC-----EGKSTLVPTICISQCLPVPANI-----GPTRI 162
Db 127 GGYETFFSYEBECVVDVPIQOKIESRALLISQCK---PVNVSVRPAYDQGGVEI 182
QY 163 LPNLYGCGQDVLNKLIIQNGIGVYVYDSCPKDPFIPESHFLRPVNDSCFKILPW 222
Db 183 LPFLYIGSAVYASKCEFLANLITLVNVSRRTSEA-CMTHLHYKWIPEVDSTHAISSH 241
QY 223 LDKSDVFIKAVANGCYLVHCLAGISRSATTAIAYIMKMDSLDEAYRFVYKERTSP 282
Db 242 FOEALIDPDCVREKGVLVHCEAGISRSPTICMAIYMKTRKQFRLKAPDIYIKORSSWS 301
QY 283 PNENFLQDLYEKIKINOT 302
Db 302 PNFGFMQQLQYESEILPST 321
```

RESULT 10

T64405

hypothetical protein DKFp43401321.1 - human
C/Species: Homo sapiens (man)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #ext_change 21-Jun-2002
C/Accession: T64405
R./Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, U.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000

A:Reference number: 223034
 A:Accession: T46405
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-303 <AAA>
 A:Cross-references: EMBL:AL137704
 A:Experimental source: adult testis; clone DKFZp43401321
 C:Genetics:
 A:Note: DKFZp43401321.1

Query Match 20.4%; Score 319; DB 2; Length 303;
 Best Local Similarity 37.7%; Pred. No. 6.5e-19;
 Matches 72; Conservative 33; Mismatches 70; Indels 16; Gaps 5;

QY 123 GGFPEFRCPGCGEGSTL-----VPTCSQ-----C-LPVAINT-GPTITLNLX 167
 DB 54 GGYRFRSEYEFCSKTKALAIAPPVPSPATETPLDGCSCGCPHLDGQGPVILPFLY 113
 QY 168 LGCCRDVNLKEILQONGIGYVLANASYTCPPKDPFIPESHFLRPVNDSECEKILPMLDKSV 227
 DB 114 LGSYVHAARDMDALGITALLNVSDCPN-HFEGHQYKCIPEVDNKKADISSWEMAI 172
 QY 228 DFEKAKASNGCVLVHCLAGISRSATIAIAYIMKMDMSIDEAYRFYKERTPTSPNPF 287
 DB 173 EYIDAVKDCRGVLVHCOAGISRSATICLAYIMKKKVRLEAFEPYKORRSIISPNFSF 232
 QY 288 LGQLLDYEKX 298
 DB 233 MGQLQFESQV 243

RESULT 11

hypothetical protein C05B10.1 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C/Accession: T32494
 R:Geisler, C.; Mameley, P.
 Submitted to the EMBL Data Library, December 1997
 A:Description: The sequence of C. elegans cosmid C05B10.
 A:Reference number: 221178
 A:Accession: T32494
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-365 <GEI>
 A:Cross-references: EMBL:AF036685; PIDN:AA88308.1; GSPDB:GN00022; CESP:C05B10.1
 A:Experimental source: strain Bristol N2; clone C05B10
 C:Genetics:
 A:Gene: CESP:C05B10.1
 A:Map position: 4
 A:Introns: 22/2; 46/3; 163/2; 204/2; 248/1; 287/2

Query Match 18.5%; Score 289; DB 2; Length 365;
 Best Local Similarity 24.9%; Pred. No. 2.5e-16;
 Matches 80; Conservative 67; Mismatches 134; Indels 40; Gaps 7;

QY 7 GTQVTRVALLESGETKVLIDSRFPVEYNTSHIEAININCSKLMKRLQODKVLIT 66
 DB 12 GBEISAQFNRIFHE--RNYIVLDCRS---NGDSVRAVRLRLPALLORLMGSMRLS 65
 QY 67 ELIOHSAKHVVDIDCSQKVVVYDQSSD-----VASISDCLVTVLGKLEKSPSVH 119
 DB 66 TVPRLKDLNLSPPDCPEVLILPGDSEDBQISAAALANLSNHRHFLVLSPVYTLISQF 125
 QY 120 LLAGGFAEFS-----RCFPGLCBGKSTLVPTCISQPCLVANIG----- 158
 DB 126 PTLRADADEMNTTFQNMNMPGASGQ-----ASSGPILNLNLRLLEGDQGGKQRAEF 180
 QY 159 PTRLPLNLVYGCQDVNLKEILQONGIGYVLANASYTCPPK-DPIFESHFLRPVNDSECE 217
 DB 181 PVKLTNLFYLGNETAKNRDVLTKKYSISHVINVTSNLPNTFEEDPNNRKYRISADDNASH 240
 QY 218 KILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSATIAIAYIMKMDMSIDEAYRFYKEX 277

DB 241 NLTKRFEPAISFIDDAARRNDSACLVHCLAGISRSVITCLAYIMKTEKCTDSAYEAWQKR 300
 QY 278 RPTISPNFNLGQLDYEKXI 298
 DB 301 NASIAPIHFHMQLTDYEKML 321

RESULT 12

T39698
 protein tyrosine phosphatase - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C/Accession: T39698
 R:Wood, V.; Skelton, J.; Churcher, C.M.; Rajandream, M.A.; Barrell, B.G.
 Submitted to the EMBL Data Library, July 1999
 A:Reference number: 221870
 A:Accession: T39698
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-330 <MOO>
 A:Cross-references: EMBL:AL109652; PIDN:CAB51765.1; GSPDB:GN00067
 A:Experimental source: strain 972h-; cosmid c17A3
 C:Genetics:
 A:Gene: p1041
 A:Map position: 2

Query Match 14.3%; Score 223.5; DB 2; Length 330;
 Best Local Similarity 39.3%; Pred. No. 5.9e-11;
 Matches 53; Conservative 18; Mismatches 63; Indels 1; Gaps 1;

QY 160 TRILPMLYLGCQDVNLKEILQONGIGYVLANASYTCPPKDPFIFESHFLRPVNDSECEKXI 219
 DB 48 SEISKLYLSWKTASELVSQKIDYTLISAMSNPNLS-VBQOHLWLQIEDSSQNT 106
 QY 220 LPMLDKSVDFIEKAKASNGCVLVHCLAGISRSATIAIAYIMKMDMSIDEAYRFYKEXRP 279
 DB 107 LQYFEKSNKRTIARLSKNNAVLVHCFRAGISRSYTVAAIYIMKMDMSIDEAYRFYKEXRP 166
 QY 280 TISPNFNLGQLDYE 294
 DB 167 GISPNANFLRLQLRY 181

RESULT 13

hypothetical protein F13D11.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001
 C/Accession: T16056
 R:Fullon, L.
 Submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid F13D11.
 A:Reference number: S69020
 A:Accession: T16056
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-186 <FUL>
 A:Cross-references: EMBL:U40939; NID:g1072175; PID:g1072176; PIDN:AAA81700.1; CESP:F13D11
 C:Genetics:
 A:Gene: CESP:F13D11.3
 A:Introns: 30/3; 57/3; 85/3; 125/3; 172/3

Query Match 14.2%; Score 222.5; DB 2; Length 186;
 Best Local Similarity 36.2%; Pred. No. 3.4e-11;
 Matches 51; Conservative 31; Mismatches 50; Indels 9; Gaps 3;

QY 160 TRILPMLYL---GCQDVNLKEILQONGIGYVLANASYTCPPKDPFIFESHFLRPVNDSECE 216
 DB 12 TVRPHFLFLAGYGC---TTPSLKQYNITHGVDCNLKTKP--IKGLDRIEVPVDNLT 65
 QY 217 EKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSATIAIAYIMKMDMSIDEAYRFYKEX 276

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:13:37 ; Search time 6.55843 Seconds
(without alignments)
2397.707 Million cell updates/sec

Title: US-10-029-345a-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMTGTQVTERIVALLLE.....PNEFPLGQIDYERKIKNQRT 302

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1552	99.2	665	1	DUSG_HUMAN
2	988	63.2	625	1	DUS8_HUMAN
3	985	63.0	663	1	DUS8_MOUSE
4	469	30.0	482	1	DUS8_HUMAN
5	468	29.9	483	1	DUS8_MOUSE
6	467.5	29.9	381	1	DUS6_HUMAN
7	462.5	29.6	381	1	DUS6_MOUSE
8	461.5	29.5	381	1	DUS6_MOUSE
9	450	28.8	619	1	VHPI_CAEEL
10	431	27.6	375	1	DUS4_CHICK
11	425	27.2	367	1	DUS1_MOUSE
12	425	27.2	394	1	DUS4_HUMAN
13	423.5	27.1	395	1	DUS4_MOUSE
14	418	26.7	314	1	DUS2_HUMAN
15	415	26.5	367	1	DUS1_HUMAN
16	415	26.5	367	1	DUS1_MOUSE
17	409	26.2	318	1	DUS2_MOUSE
18	408	26.1	384	1	DUS9_HUMAN
19	385.5	24.6	320	1	DUS7_HUMAN
20	382.5	24.5	320	1	DUS7_MOUSE
21	358.5	22.9	384	1	DUS5_MOUSE
22	357.5	22.9	280	1	DUS7_MOUSE
23	351	22.4	384	1	DUS5_HUMAN
24	262	16.8	198	1	DUSE_HUMAN
25	261	16.7	198	1	DUSE_MOUSE
26	205.5	13.1	188	1	DUSI_HUMAN
27	205.5	13.1	217	1	DUSJ_HUMAN
28	205	13.1	223	1	STRY_MOUSE
29	204	13.0	223	1	STRY_HUMAN
30	202	12.9	364	1	PVHI_YEAST
31	201.5	12.9	313	1	MSY_HUMAN
32	195.5	12.5	185	1	DUS3_HUMAN
33	194	12.4	185	1	DUS3_MOUSE

34	185	11.8	807	1	YB9T_YEAST
35	183.5	11.7	276	1	PTR3_CHLEU
36	181.5	11.6	278	1	PMPI_SCHPO
37	178.5	11.4	295	1	DUSF_HUMAN
38	177	11.3	125	1	DUSF_MOUSE
39	171.5	11.0	339	1	DUSC_MOUSE
40	171	11.0	339	1	MSG5_YEAST
41	158.5	10.9	340	1	DUSC_HUMAN
42	157	10.9	209	1	YI13_YEAST
43	155	9.9	198	1	DUSD_HUMAN
44	152	9.7	292	1	Y042_CAEEL
45	150	9.6	198	1	DUSD_MOUSE

ALIGNMENTS

RESULT 1
DUSG_HUMAN STANDARD; PRT; 665 AA.
AC Q9BY64; Q9COG3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase
DE phosphatase 7) (MKP-7).
GN DUSP16 OR MKP7 OR KIAA1700.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,
RT functions as a shuttle protein".
RL J. Biol. Chem. 276:39002-39011(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from Brain which code
RT for large proteins in vitro".
RL DNA Res. 7:347-355(2000).
CC -1- FUNCTION: Involved in the inactivation of MAP kinases.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL: AB052156; BAB40814.1; -;
CC EMBL: AB051487; BAB21791.1; ALT_INIT.
CC HSPSP; Q16828; MKP.
CC Genew; HGNC:17909; DUSP16.
CC MIM: 607175; -;
CC GO: GO:0005737; C:cytoplasm; TAS.
CC GO: GO:0005634; C:nucleus; TAS.

```
DR GO; GO:0004721; F:protein phosphatase activity; TAS.
DR GO; GO:0016311; P:dephosphorylation; TAS.
DR GO; GO:0000188; P:inactivation of MAPK; TAS.
DR GO; GO:0045209; P:pleiotomycin B-sensitive MAPK phosphatase nuc. .; TAS.
DR GO; GO:0045204; P:MAPK nucleus export; TAS.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase; Nuclease protein.
FT DOMAIN 22 137 RHODANES.
FT DOMAIN 227 289 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 244 244 PHOSPHOCYSTEINE INTERMEDIATE (BY
SQ SEQUENCE 665 AA; 73101 MW; 1BD853FF08460DF CRC64;
SIMILARITY)
```

```
Query Match 99.2%; Score 1552; DB 1; Length 665;
Best Local Similarity 99.3%; Pred. No. 2.5e-125;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MAHEMIGTQIVTERLVALLSSTGTEKVLIDSRPFVENVSHILEANINCSKLMKRRLOQ 60
DB 1 MAHEMIGTQIVTERLVALLSSTGTEKVLIDSRPFVENVSHILEANINCSKLMKRRLOQ 60
QY 61 DKVLITLILHSAKHKVDIDCSQKVVVYDQSSQDVASLSSDCLFTVLGLKLEKSFNSVHL 120
DB 61 DKVLITLILHSAKHKVDIDCSQKVVVYDQSSQDVASLSSDCLFTVLGLKLEKSFNSVHL 120
QY 121 LAGGFAPFRCRPGFLCEGKSTVPTCTISOPCLPVANIGTRLLPLVYIGCQRDVANKELI 180
DB 121 LAGGFAPFRCRPGFLCEGKSTVPTCTISOPCLPVANIGTRLLPLVYIGCQRDVANKELI 180
QY 121 LAGGFAPFRCRPGFLCEGKSTVPTCTISOPCLPVANIGTRLLPLVYIGCQRDVANKELI 180
DB 121 LAGGFAPFRCRPGFLCEGKSTVPTCTISOPCLPVANIGTRLLPLVYIGCQRDVANKELI 180
QY 181 QONGIGYVLNASTYCPKDPFIPESHFLRPVNDSCFCEKILPWLDSVDFIKAKASNGCV 240
DB 181 QONGIGYVLNASTYCPKDPFIPESHFLRPVNDSCFCEKILPWLDSVDFIKAKASNGCV 240
QY 241 IVHCLAGISRSATIAIAYIMKMDMSLDEAARFVKEKRPITSPNPNFLGQLLDYEKKIKN 300
DB 241 IVHCLAGISRSATIAIAYIMKMDMSLDEAARFVKEKRPITSPNPNFLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
```

```
RESULT 2
DUS8_HUMAN STANDARD; PRT; 625 AA.
```

```
AC 013202;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Dual specificity protein phosphatase hvh-5).
GN DUSP8 OR VH5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCI1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96009533; PubMed=7561881;
RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.B.;
```

```
RT "hvh-5: a protein tyrosine phosphatase abundant in brain that
RT inactivates mitogen-activated protein kinase.";
RL J. Neurochem. 65:1823-1833(1995).
CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
CC phosphate as well as with serine/threonine-protein phosphate (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- TISSUE SPECIFICITY: Abundant in brain, heart and skeletal muscle.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; U27193; AAA83151.1; -
DR HSSP; Q16828; IMKP
DR Genew; HGNC:3074; DUSP8.
DR MIM; 602038; -
```

```
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.
DR GO; GO:0000188; P:inactivation of MAPK; TAS.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase; Nuclease protein.
FT DOMAIN 23 138 RHODANES.
FT DOMAIN 162 430 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 310 550 PRO-RICH.
FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
SQ SEQUENCE 625 AA; 65840 MW; DCBEA1487219666 CRC64;
SIMILARITY)
```

```
Query Match 63.2%; Score 988; DB 1; Length 625;
Best Local Similarity 63.1%; Pred. No. 5.2e-77;
Matches 190; Conservative 55; Mismatches 54; Indels 2; Gaps 2;
```

```
QY 1 MAHEMIGTQIV-TERLVALLSSTGTEKVLIDSRPFVENVSHILEANINCSKLMKRRLOQ 59
DB 1 MAGDRLPRKWDADKALSLILNGRGPVLVDSRFRVYNSHNVLSVNICCSKLVKRRLOQ 60
QY 60 QDKVLITLILHSAKHKVDIDCSQKVVVYDQSSQDVASLSSDCLFTVLGLKLEKSFNSVH 119
DB 61 QDKVLITLILHSAKHKVDIDCSQKVVVYDQSSQDVASLSSDCLFTVLGLKLEKSFNSVH 119
QY 120 LLAGGFAPFRCRPGFLCEGK-STVPTCTISOPCLPVANIGTRLLPLVYIGCQRDVANKELI 178
DB 121 ILTGFPATFSSCFGLCEGKRALPLMSLSQCLPVPSVGLTILPLLYGSKQDVANKXO 180
QY 179 LIQONGIGYVLNASTYCPKDPFIPESHFLRPVNDSCFCEKILPWLDSVDFIKAKASNG 238
DB 181 LMTONGIGYVLNASTYCPKDPFIPESHFLRPVNDSCFCEKILPWLDSVDFIKAKASNG 240
QY 239 CVLVHCLAGISRSATIAIAYIMKMDMSLDEAARFVKEKRPITSPNPNFLGQLLDYEKKIKN 298
```

DB 241 QVIVHCLAGISRSATIAIAYIMKTGMSSDDAYRFVDRRRSPISPNFGLLEYERTL 300
QY 299 K 299
DB 301 K 301

RESULT 3
DUSA_MOUSE STANDARD; PRT; 663 AA.
ID DUSA_MOUSE 009112;
AC 009112;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Neutonal tyrosine threonine phosphatase 1).
GN DUSP8 OR NTP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=63111565; PubMed=8731137;
RA Theodosiou A.M., Rodriguez N.R., Nesbit M.A., Ambrose H.J.,
RA Peterson H., McJellan-Arnold B., Boyd Y., Leverina M.A., Owen N.,
RA Blake D.J., Ashworth A., Davies K.E.;
RT "A member of the MAP kinase phosphatase gene family in mouse
RT containing a complex trinucleotide repeat in the coding region."
RL Hum. Mol. Genet. 5:675-684 (1996).
CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
CC phosphate as well as with serine/threonine-protein phosphate (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- TISSUE SPECIFICITY: Expressed predominantly in brain and lung.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: X95518; CA64772.1; -.
DR HSSP: Q16828; IMKP.
DR MGD: MGI:106626; Dusp8.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR PRINTS: PR01764; MAPKPHPTASE.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANSE_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydroxase; Nuclear protein.
FT DOMAIN 23 138 RHODANSE.
FT DOMAIN 162 432 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 452 459 POLY-ARG.
FT DOMAIN 555 558 POLY-SER.
FT DOMAIN 559 576 POLY-GLY.

FT DOMAIN 577 600 POLY-SER.
FT DOMAIN 311 552 PRO-RICH.
FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SEQUENCE 663 AA; 66847 MW; 416f429a12c1fa7c CRC64;
SIMILARITY)).
SQ
Query Match 63.0%; Score 985; DB 1; Length 663;
Best Local Similarity 62.5%; Pred. No. 1e-76;
Matches 188; Conservative 55; Mismatches 56; Indels 2; Gaps 2;
QY 1 MAHEMIGTOIV-TERLVALLSGTEKVLIDSPPEVYNTSHLEAININCSKLMKRRLO 59
DB 1 MADDRLPRKVMADKXLLRGGPGPLVIDSRSPFVNSCHVLSVNICSLTKVRRLO 60
QY 60 QDVLVLTTELIGHAKKVIDCSQKVVYVDDSQDVASLSSDCEFTVTLGKLEKSFVS 119
DB 61 QGRVTAELIOPATRSQVADTEPDVVVDDQSTRDASVLAADSLSTLSKLGCCPQVA 120
QY 120 LLAGFAEFRCPCPGCEGKSTLVPT-CISQPLPVANIGPTRILPMLYLGCORVLANKE 178
DB 121 ILTNGFATRSCTPGCEGKPAFLPMSLSQCLPVPVGLTIRLPHLYGSGQKDVANKD 180
QY 179 LIOQNGIGVYVNASYTCPRKDFIPESHFLRPVNDSEFCXKILPWLKSVDFIKAKASNG 238
DB 181 LMTQNGISYVYVNASNSCPKDFICESRFRMIPINDVYCEKTLPLWLDKSIKXKSSC 240
QY 239 CVLVHCLAGISRSATIAIAYIMKTGMSSDDAYRFVDRRRSPISPNFGLLEYERTL 298
DB 241 QVIVHCLAGISRSATIAIAYIMKTGMSSDDAYRFVDRRRSPISPNFGLLEYERTL 300
QY 299 K 299
DB 301 K 301

RESULT 4
DUSA_HUMAN STANDARD; PRT; 482 AA.
ID DUSA_HUMAN 09Y6W6;
AC 09Y6W6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase
DE phosphatase 5) (MKP-5).
GN DUSP10 OR MKP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99312929; PubMed=10391943;
RA Tanoue T., Moriguchi T., Nishida E.;
RT "Molecular cloning and characterization of a novel dual specificity
RT phosphatase, MKP-5."
RL J. Biol. Chem. 274:119949-19956 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20065165; PubMed=10597297;
RA Theodosiou A., Smith A., Gillieron C., Arkinetall S., Ashworth A.;
RT "MKP5, a new member of the MAP kinase phosphatase family, which
RT selectively dephosphorylates stress-activated kinases."
RL Oncogene 18:6981-6988 (1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, Lung, and Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.B., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

QY 85 VVVVDDSSQDVASLSSDCEFTLVLLGKLEKSEFNSVHLLAGFAERSRCFPGLCGK----- 139
Db 223 IIVDENINERNSRWPSPQPLATIVESLKRGEKPELVKLGSLSPFKQHNHLCONSLOQE 292
QY 140 -----STLVPTCISQCP-P-VANIGPFRILPNVYLGGORVNLKELIQNGIGY 187
Db 293 CREVGAGSAASASLIPQPI--PTTPDIENALFLIPFLPFLNGQADIDYMQRLNIGY 350
QY 188 VLNAS-----YTCPEKPFIPESHFLEVPVNDSCCEKILWDLSDVFIEKARASNGCVLV 242
Db 351 VINTVTHLPVYHEKGLF-----NYKSLPPTDSKNQMLRGYFEEAEFFIEEAHQCGKGLLI 406
QY 243 HCLAGISRSKNTLIATVIMKRMDSUDEAARYEYKEXRPTTSPNFNFVGGOLDYIEKIKN 300
Db 407 HCGAGVSRSAATVIAVIMKMTMTMDAYKFKGRPIISPLNMGQLFEEDLN 464

RESULT 5
DUSA MOUSE
ID DUSA MOUSE STANDARD; PRT: 483 AA.
AC Q9ES50; Q9CZ9; 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase phosphatase 5) (MKP-5).
GN DUSP10 OR MKP5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20515587; Pubmed=11060451;
MA Masuda K., Shima H., Kikuchi K., Matsubae Y., Matsuda Y.;
RT "Expression and comparative chromosomal mapping of MKP-5 genes
R1 DUSP10/Dusp10".
RL Cytogenet. Cell Genet. 90:71-74(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/60; TISSUE=Embryo;
RX MEDLINE=21085660; Pubmed=11217851;
MA Kawai U., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,
RA Aizawa T., Hara A., Fukunishi Y., Komno H., Akachi U., Fukuda S.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Knehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
RA Sakai K., Steubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake U., Boiteill D., Bojunga N., Aono H., Baldarelli R., Barih G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.F.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima U., Mazzarella J., Mombauts P.,
RA Norcone P., Ring B., Ringwald M., Rodriguez I., Sakamoto K.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch C.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:695-690(2001).
CC -I- FUNCTION: Involved in the inactivation of MAP kinases. Has a
specificity for the MAPK1/MAPK2/MAPK3/MAPK4 subfamily (By
similarity).
CC -I- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
phosphate.
CC -I- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
phosphate.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
CC -I- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.

[illegible]

CN DUSP6 OR MKP3 OR PVST1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Forebrain;
 RX MEDLINE=96311959; PubMed=8670865;
 RA Groom L.A., Sneddon A.A., Aleesi D.R., Dowd S., Keyse S.M.;
 RT "Differential regulation of the MAP, SAP and RK/p38 kinases by Pyc1,
 RL a novel cytosolic dual-specificity phosphatase.";
 RN EMO J. 15:3621-3632(1996).
 RP [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Liver;
 RX MEDLINE=99077745; PubMed=9858808;
 RA Furukawa T., Yatsunaka T., Yousef E.M., Abe T., Yokoyama T.,
 RA Fukuishi S., Soeda E., Hoshi M., Hayashi Y., Sunamura M., Kobai M.,
 RA Horii A.;
 RT "Genomic analysis of DUSP6, a dual specificity MAP kinase phosphatase,
 RT in pancreatic cancer.";
 RL Cytochrome. Cell Genet. 82:156-159(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT VAL-114.
 RC TISSUE=Colon, Kidney, Skin, and Stomach;
 RX MEDLINE=22388257; PubMed=12477933;
 RA Krausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Stausberg R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Dichtenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Steplonen M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Roark S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunatratte P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Scherch A., Schein J.B., Jones S.J.W., Maira M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 204-347.
 RX MEDLINE=99140299; PubMed=10048930;
 RA Stewart A.E., Dowd S., Keyse S.M., McDonald N.Q.;
 RT "Crystal structure of the MAPK phosphatase Pyc1 catalytic domain and
 RT implications for regulated activation";
 RL Nat. Struct. Biol. 6:174-181(1999).
 CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
 CC family.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q16828-1; Sequence=Displayed;
 CC Name=2; Synonyms=DUSP6-ALT;
 CC IsoId=Q16828-2; Sequence=VSP_005137;
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carrincci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Feller J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalek U., Smallue D.E.,
RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
CC family (By similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AK005062; BAB23786.1; -
DR EMBL: AK009131; BAB26093.1; -
DR EMBL: BC003869; AA003869.1; -
DR HSSP: Q16828; IMKP.
DR MGD: MG1:1914853; Dusp6.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR PRINTS: PR01764; MAPKPHPTASE.
DR SMART: SM00195; DSPC; 1.
DR SMART: SMO0450; RHOD; 1.
DR PROSITE: PSS0206; RHODANSE_3; 1.
DR PROSITE: PSS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE: PSS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PSS50054; TYR_PHOSPHATASE_DUAL; 1.
DR KMW: Hydrolyase.
FT DOMAIN 30 148 RHODANSE.
FT DOMAIN 206 381 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 293 293 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
FT CONFLICT 22 22 W -> G (IN REF. 1; BAB26093).
FT CONFLICT 34 34 L -> F (IN REF. 1; BAB26093).
SQ SEQUENCE 381 AA; 42407 MW; 7EALPB154FAD2DA CRC64;

Query Match 29.6%; Score 462.5; DB 1; Length 381;
Best Local Similarity 33.7%; Pred. No. 3.2e-32;
Matches 109; Conservative 64; Mismatches 105; Indels 45; Gaps 7;

Qy 19 LESGTEKVLILDSRPFEVNTSHILEININCSKLMKRLQODKVLITELIQH-SAKHKV 77
Db 27 LEIGNEHLLMDCRPGQLYESSHESINVAIEGIMLRRLQKGNLPRALFTCEPDRDF 86
Qy 78 DIDC-SQKVVVYQSSQD-VASLSSDCLTVLLGKLEKSFNSVHLLAGFAEFSRCPGL 135
Db 87 TRCGGTFTVVLVDENSDMNMENGTGSEVLGLLLKLLKDECCRAFLVLEGFSKFOAEPAH 146

Qy 136 CEKSTLVPTC-ISOCPCLPVANIG----- 158
Db 147 CE--TWLDGSSSSPPLPLVGLGGRISRSDSSDIESDLDRPNRGTSDGSPLSNQ 204
Qy 159 --PTRIPLNYLGCORVDINKELIQNGIGVYNASVYCPKP-DFTPESHFLRPVNSF 215
Db 205 SFPEVLPLFLYGCAQDSTLDVLEFGIKYILNTPNLPNLENAGDFPKQIPISDHW 264
Qy 216 CEKILWLDKSDVDFEKARASNCVYHCLAGISRSATLAIYIMKMDSLDEARFVK 275
Db 265 SQNLISQFFPEALSFIDEARKNGVHVCILAGISRSVTVAVYIMOKNLISMDADIVK 324
Qy 276 EKRPISPNPFLGOLLDYEKKI 298
Db 325 MKSNISPNPFWGLDPERTL 347

RESULT 8
DUSE_RAT
ID DUSE_RAT STANDARD; PRT; 381 AA.
AC 064346;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
DE phosphatase 3) (MKP-3).
GN DUSP6 OR MKP3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NBL_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Neuron;
RX MEDLINE=96216487; PubMed=8631996;
RA Moudy R.J., Vega Q.C., Campbell J.S., Wenderoth M.P., Hauschka S.D.,
RA Krebs E.G., Dixon J.E.;
RT "A novel cytoplasmic dual specificity protein tyrosine phosphatase
RT implicated in muscle and neuronal differentiation."
RL J. Biol. Chem. 271:3795-3802(1996).
CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
CC family. Implicated in muscle and neuronal differentiation.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X94185; CAA63895.1; -
DR EMBL: U42627; AAB06202.1; -
DR HSSP: Q16828; IMKP.
DR InterPro: IPR000340; DS_phosphatase.

Query Match	29.5%	Score 461.5;	DB 1;	Length 381;
Best Local Similarity	33.7%	Pred. No. 3.9e-32;		
Matches 109; Conservative	64.1	Missed 10		

RESULT 9

DR EMBL, U23178; AAC46719.1; -.
DR PIR; T15969; T15969.
DR HSSP: O16828: 1MKD

Query Match	28.8%	Score 450;	DB 1;	Length 619;
Best Local Similarity	50.6%	Pred. No. 6.9e-31;		
Matches	90;	Conservative	34;	Mismatch 1

RESULT 10

```

CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF167296; AAD46656.1; -.
CC HSSP: Q16828; IMKP.
CC InterPro: IPR000340; DS phosphatase.
CC InterPro: IPR001763; Rhodanese-like.
CC InterPro: IPR000387; TYR_phosphatase.
CC Pfam: PF00782; DSPC; 1.
CC Pfam: PF00581; Rhodanese; 1.
CC PRINTS: PR01764; MAPKPHPTASE.
CC SMART: SM00195; DSPC; 1.
CC SMART: SM00450; RHOD; 1.
CC PROSITE: PS50206; RHODANESE_3; 1.
CC PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
CC PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
CC PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
CC HydroLase; Nuclear protein.
CC DOMAIN 25 143 RHODANESE.
CC FT DOMAIN 178 375 PROTEIN-TYROSINE PHOSPHATASE.
CC FT ACT_SITE 261 261 PHOSPHOCYSTEINE INTERMEDIATE (BY
CC SIMILARITY).
CC SO SEQUENCE 375 AA; 41052 MW; 179290DC2BEEF1 CRC64;
Query Match 27.6%; Score 431; DB 1; Length 375;
Best Local Similarity 33.2%; Pred. No. 1.6e-29;
Matches 104; Conservative 65; Mismatches 120; Indels 24; Gaps 9;
QY 4 EMIGTOIVTERLVALLSEGTETKVLIDSRPFVEYNTSHLEATININCKLMKRRLOQDKV 63
DB 9 EMGGSAL--RLVGRERASGRCILDCRFPLASAHITGALNVRNCTIVRRR-AKGAV 65
QY 64 LITELI--QHSARKKVDIDCSQKVVYDQSSODVASLSDCFLVLGLKEGF--NSVH 119
DB 66 SLRGQILPAEGBVBARLAAGLYTNVLYLDESRPAELALRDSTVALVRLARRMARADIR 125
QY 120 LLAGGFAPESRCPPGLCEGKSTLVPTGISDP-----C-LPVANI-GPTRILEN 165
DB 126 LLAGGYRFRFASVEYFPCAKTKTL--SSISPPSSAESLIDLGSSCGTPLHDQGGVELLPF 183
QY 166 LVYGCQGVAVLKEILQONGIGVYNASVYTCPPKDPFIPESHPLRVAVVDSFCCKLPLMDK 225
DB 184 LVYGSAYHAARRDMLDALGITLALNVSSDCPN-HFEGHYQYKCIPEVDNKAADISSPFME 242
QY 226 SVDFIEKAKASNGCVLVHCLAGISRSATIAIAYIMKMDSLDEAVRFVVEKEPPTSPNF 285
DB 243 AIEYIDSVKCCGCVLVHCGAGISRSATICLAYIMMKRYLKEAFAEFVQGRSIIISPNF 302
QY 286 NPLGQLLDYEKKI 298
DB 303 SFMGQLQFESQV 315

```

```

DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE 3CH134) (Protein-tyrosine phosphatase ERP).
GN DUSP1 OR PTPN10 OR MKP1 OR 3CH134 OR PTPN16.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=92158357; PubMed=1741163;
RA Charles C.H., Abler A.S., Lau L.F.;
RT "cDNA sequence of a growth factor-inducible immediate early gene and
RT characterization of its encoded protein."
RL Oncogene 7:187-190(1992).
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93360956; PubMed=8355678;
RA Noguchi T., Metz R., Chen L., Mattei M.-G., Carrasco D., Bravo R.;
RT "Structure, mapping, and expression of erp, a growth factor-inducible
RT gene encoding a nontransmembrane protein tyrosine phosphatase, and
RT effect of ERP on cell growth."
RL Mol. Cell. Biol. 13:5195-5205(1993).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary gland;
RX MEDLINE=92388257; PubMed=12477932;
RA Stranberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schur G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Heien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
RA Brownstein M.J., Uedon T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.,
RA Villalon D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rodierfeld Y.S.N., Krzywinski M.I., Skalska U., Smilans D.E.,
RA Schnerch A., Schen J.E., Jones S.J.M., Maria W.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN CHARACTERIZATION.
RX MEDLINE=94037096; PubMed=8221888;
RA Sun H., Charles C.H., Lau L.F., Tonks N.K.;
RT "MKP-1 (3CH134), an immediate early gene product, is a dual
RT specificity phosphatase that dephosphorylates MAP kinase in vivo."
RL Cell 75:487-493(1993).
RN [5]
RN FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
RN kinase ERK2 on both Thr-183 and Tyr-185.
RN CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
RN tyrosine + phosphate.
RN CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
RN phosphate.
RN INDUCTION: By growth factors.
RN SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
RN Non-receptor class dual specificity subfamily.
RN SIMILARITY: Contains 1 rhodanese domain.
RN -----
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation-
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to license@isb-sib.ch.

```

CC -----
CC EMBL, X61940; CAA43944.1; -
CC DR EMBL, S64851; AAB27882.1; -
CC DR EMBL, BC006967; AAH06967.1; -
CC DR PIR, A54681; S24411.
CC DR HSSP, Q16828; IMKP.
CC DR MGJ, MG1:105120; Dusp1.
CC DR InterPro; IPR000340; DS phosphatase.
CC DR InterPro; IPR008343; MAPK phosph.
CC DR InterPro; IPR001763; Rhodanese-like.
CC DR InterPro; IPR000387; TYR_phosphatase.
CC DR Pfam; PF00782; DSPC; 1.
CC DR Pfam; PF00581; Rhodanese; 1.
CC DR PRINTS; FRO1764; MAPKPHPTASE.
CC DR SMART; SM00450; RHOD; 1.
CC DR SMART; SM00195; DSPC; 1.
CC DR PROSITE; PSS0206; RHODANES_3; 1.
CC DR PROSITE; PSS0383; TYR_PHOSPATASE_1; 1.
CC DR PROSITE; PSS0056; TYR_PHOSPATASE_2; 1.
CC DR PROSITE; PSS0054; TYR_PHOSPATASE_DUAL; 1.
CC DR Hydrolase; Cell cycle.
CC FT DOMAIN 20 137 RHODANES.
CC FT ACT_SITE 175 367 PROTEIN-TYROSINE PHOSPHATASE.
CC FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE.
CC FT MUTAGEN 258 C->S; LOSS OF ACTIVITY.
CC SQ SEQUENCE 367 AA; 39369 MW; 50B5F90FEBBD19AB CRC64;

```

Query Match Best Local Similarity 27.2%; Score 425; DB 1; Length 367; Matches 104; Conservative 48; Mismatches 129; Indels 22; Gaps 5;

```

QY 15 LVALDSGTEKVLIDSPFVNTSHILRLININCKMKRLQODKVLTELIOHSAK 74
DB 13 LRLALREBAACCLLDGCRSPFAFAGHIAQVAVRSTIVRRRAKAMGLEHIVNAELR 72
QY 75 HKVDIDGQKVVYVDGSSQDVASISDPCFLTVLGLK--EKSPNSVHLLAGFAFESKCF 132
DB 73 GRLLAGVYHAAVLLDERSASLIDGAKRDGTLAAGLACHEASTQVFTLQGGYEAFFASAC 132
QY 133 PGLCEGKSTLVPTCTISOP-----CLPVAANI-----GPTRIIPNLVIGCCORDVL 175
DB 133 PELCSKST--PTCLSPILSTSPDPAESGSCSSCPPLVDQGPPEILSPFLYLSGAVHAS 190
QY 176 NKELIQNGIGYVLNASYTCPEKDFIPESHPLRVPNDSFCERKILPMLDKSVPTERAKA 235
DB 191 RKMLDALGITALINYSANCPN-HFEGHYQKSIIVEDNHKADISSWFEKADIDFISIKD 249
QY 236 SNGCVLVHCLAGISRSATTAIYIMKRMMSLDEAYRVTVEKERPTISNENFQLQLDYE 295
DB 250 AGGRVYVHQAQGSRSATTCIAIYIMKRVKVLDEAFEFVKORRSIISPNFSFGQLQF 309
QY 296 KKI 298
DB 310 SQV 312

```

RESULT 12
DUSA HUMAN
ID DUSA 01315; 013524; STANDARD; PRT; 394 AA.
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase phosphatase-2) (MKP-2) (Dual specificity protein phosphatase hvm2).
GN DUSP4 OR MKP2 OR VR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxId=9606;
RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=95221370; PubMed=7535768;
RT Guan K.-L., Butch E.;
RT "Isolation and characterization of a novel dual specific phosphatase,
RT HVM2, which selectively dephosphorylates the mitogen-activated
RT protein kinase.";
RL J. Biol. Chem. 270:7197-7203 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96198119; PubMed=8626452;
RT Chu Y., Solis P.A., Khosravi-Far R., Der C.J., Kelly K.;
RT "The mitogen-activated protein kinase phosphatases PAC1, MKP-1, and
RT MKP-2 have unique substrate specificities and reduced activity in vivo
RT toward the ERK2 sevenmarer mutation.";
RL J. Biol. Chem. 271:6497-6501 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin, and Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Dietzenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Cavaant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton B., Ketterman M., Madan A., Rodriguez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerker A., Schein J.B., Jones S.J.M., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases
CC ERK1 and ERK2.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, U21108; AAA65119.1; -
CC DR EMBL, U48807; AAC50452.1; -
CC DR EMBL, BC002671; AAH02671.1; -
CC DR EMBL, BC014565; AAH14565.1; -
CC DR HSSP, Q16828; IMKP.
CC DR Genew; HGNC:3070; DUSP4.
CC DR MIM; 602747; -
CC DR GO; GO:0005634; C:nucleus; TAS.
CC DR GO; GO:0008330; P:protein tyrosine/threonine phosphatase acti. .; TAS.
CC DR GO; GO:0001655; P:MAPKK cascade; TAS.
CC DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
CC DR InterPro; IPR000340; DS phosphatase.
CC DR InterPro; IPR008343; MAPK phosph.
CC DR InterPro; IPR001763; Rhodanese-like.

```

```

DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; Nuclear protein.
FT DOMAIN 41 159 RHODANASE.
FT DOMAIN 197 394 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 280 280 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT ACT_SITE 280 280 SIMILARITY).
FT CONFLICT 111 111 R -> G (IN REF. 2).
SQ SEQUENCE 394 AA; 42953 MW; 0603971759B952E CRC64;

Query Match 27.2%; Score 425; DB 1; Length 394;
Best Local Similarity 33.4%; Pred. No. 5.5e-29;
Matches 100; Conservative 64; Mismatches 113; Indels 22; Gaps 9;

QY 19 LBSGTEKVLIDSRPVEYNTSHLEAININCKMKRRLQODKVLITELI--QHSARK 76
DB 39 LPESG-GKCLLDRCPLFASAGYILGSVNVRCNTIYRR-AKGSVSLQILPAEEVRR 96
QY 77 VDDICGQKVVYVQSSQDVASLSSDCLFVLLGKLEKSF--NSVHLAGFAFSPRCFPG 134
DB 97 LRSGLYSAVIVYDERSPRASLSREDSTVSLVQALRNARTRDILCKGGERFSSEYPE 156
QY 135 LCEGKSTL-----VPTCISQP-----C-LPVANI-GPTRIILPNLYLGCQRDVINKEL 179
DB 157 FCGKTKALAIIPRPVPSPATEPDLGSSCGTPLHDGCGVELLPFLYLSAHAARRDM 216
QY 180 IQQNGIGVYVNASYTCRKPDPFIPESHFLRVVNDSECEKILPWLDSVDFIEKAKSNGC 239
DB 217 LDALGITALLNVSSDCPN-HFECHGYQKCIPEVDNHNKADISSWFEAIEYIDAVKCRGR 275
QY 240 VLVHCLAGISRATIAAYIMKMDMSLDEAYRFVVEKRTISPNRFLGQLLDYEKI 298
DB 276 VLVHCQAGISRSATICLAYIMMKRVLEAFEFVKORRSIIISPNFSFMQLQFBSQV 334

RESULT 13
DUS4_RAT STANDARD; PRT; 395 AA.
ID DUS4_RAT 062767;
AC 062767;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (Map kinase
DE phosphatase-2) (MKP-2).
GN DUSP4 OR MKP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Phenochromocytoma;
RX MEDLINE=95301550; Pubmed=7782322;
RT "A novel mitogen-activated protein kinase phosphatase. Structure,
RT expression, and regulation."
RL J. Biol. Chem. 270:14587-14596(1995).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
CC and ERK2 (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.

```

```

CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC TISSUE SPECIFICITY: Expressed at moderate levels in nearly all
CC tissues and cells including brain, spleen, and testes with the
CC higher expression in the heart and lung and lower expression in
CC skeletal muscle and kidney. Undetectable in liver. Expressed in
CC hippocampus, piriform cortex, and the suprachiasmatic nucleus.
CC -1- INDUCTION: By mitogens and by stress.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23438; AAC52493.1; -.
DR HSSP; Q16828; IMKP.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; Nuclear protein.
FT DOMAIN 42 160 RHODANASE.
FT DOMAIN 198 395 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 281 281 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT ACT_SITE 281 281 SIMILARITY).
SQ SEQUENCE 395 AA; 43187 MW; A90BFFD378A050FD CRC64;

Query Match 27.1%; Score 423.5; DB 1; Length 395;
Best Local Similarity 32.2%; Pred. No. 7.4e-29;
Matches 98; Conservative 61; Mismatches 118; Indels 27; Gaps 7;

QY 16 VALLESGETEKVLLIDSRPVEYNTSHLEAININCKMKRRLQODKVLITELI--QHSAR 73
DB 38 LGLSLSG--KCLLDRCPLFASAGYIRGSVNVRCNTIYRR-AKGSVSLQILPAEEV 94
QY 74 KHKVDIDCGQKVVYVQSSQDVASLSSDCLFVLLGKLEKSF--NSVHLAGFAFSPRC 131
DB 95 RRLRLSGLYSAVIVYDERSPRASLSREDSTVSLVQALRNARTRDILCKGGERFSSE 154
QY 132 PFGELGKSTLPTICISQCLPVANI-----GPTRIILPNLYLGCQRDV 174
DB 155 YEFSCSKTAL--AALPPVPSPSTNESLDLGCSSCGTPLHDGCGPEILPFLYLSAHTA 212
QY 175 LNKELIQQNGIGVYVNASYTCRKPDPFIPESHFLRVVNDSECEKILPWLDSVDFIEKAK 234
DB 213 AARDMDALGITALLNVSSDCPN-HFECHGYQKCIPEVDNHNKADISSWFEAIEYIDAVK 271
QY 235 ASNGCYLVHCLAGISRATIAAYIMKMDMSLDEAYRFVVEKRTISPNRFLGQLLDY 294
DB 272 DCRGRVLVHCQAGISRSATICLAYIMMKRVLEAFEFVKORRSIIISPNFSFMQLQF 331
QY 295 EKKI 298
DB 332 ESQV 335

RESULT 14
DUS2_HUMAN

```

ID DUS2_HUMAN STANDARD; PRT; 314 AA.
 AC Q05923;
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dual specificity protein phosphatase 2 (EC 3.1.3.48) (EC 3.1.3.16)
 DE Dual specificity protein phosphatase PAC-1.
 GN DUSP2 OR PAC1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93206122; PubMed=7681221;
 RA Rohan P., Davis P., Moskaluk C.A., Kearns M., Krutzsch H.,
 RA Siebenlist U., Kelly K.,
 RT "PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.";
 RL Science 259:1763-1766 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96070437; PubMed=7590752;
 RA Yi H., Morton C.C., Weremowicz S., McBride O.W., Kelly K.,
 RT "Genomic organization and chromosomal localization of the DUSP2 gene,
 RT encoding a MAP kinase phosphatase, to human Zp11.2-q11.";
 RL Genomics 28:92-96 (1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=B-cell;
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schaller G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Spletton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.C.,
 RA Rata S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: Regulates mitogenic signal transduction by
 CC dephosphorylating both Thr and Tyr residues on MAP kinases
 CC ERK1 and ERK2.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: In hematopoietic tissues.
 CC -1- INDUCTION: By mitogens.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; L11329; AAA50779.1; -.

DR EMBL; U23853; AAA6112.1; -.
 DR EMBL; BC007771; AAA6777.1; -.
 DR PIR; A57126; A57126.
 DR PIR; 11KZ2; 30-MAR-02.
 DR Genew; HGNC:3068; DUSP2.
 DR MIM; 603068; -.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.
 DR GO; GO:0008330; F:protein tyrosine/threonine phosphatase acti. .; TAS.
 DR GO; GO:0000188; P:inactivation of MAPK; TAS.
 DR GO; GO:0006478; P:protein amino acid dephosphorylation; TAS.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHRASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANASE 3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE 1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE 2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR Hydrolase; Nuclear protein; 3D-structure.
 FT DOMAIN 23 144 RHODANASE.
 FT DOMAIN 237 302 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 257 257 PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT SIMILARITY).
 SQ SEQUENCE 314 AA; 34399 MW; F03543C6B10CA5 CRC64;
 Query Match 26.7%; Score 418; DB 1; Length 314;
 Best Local Similarity 36.0%; Pred. No. 1.6e-28;
 Matches 109; Conservative 45; Mismatches 105; Indels 44; Gaps 8;
 QY 24 EKVLLIDSRPFVEYNTSHILEININCSKLMKR-----IQDDKLTLELQHS 72
 DB 25 ERTLLIDCRPLAFGRHVRARPYPMWALLRRRARGPAVLAACLPDRRLRLTVRG 84
 QY 73 AAKKVDIDSCQVYVYDSSQDVASLSD---CELVYLGKLRKFSNVYLLAGFAEF 128
 DB 85 L-----AAVVLIDSGASVVELRPDSPAHYLALHETRAGPLAVYFLRGDGF 135
 QY 129 SRCFPGLC-EGKSTLVP-----TGISQCLPVANI-GPTRLPLVYGCQDVLNKLQ 161
 DB 136 QCCPCDLCSEAPAPALPTGKTRSDSRAPVYDGGVELLPYLFGLGSCSHSSDLQGIQ 195
 QY 182 QNGIGVYVNAAGTCTCPKDPFIPESHF-----LRVPVDSFCEKILPWLDKSDVFIEKAKA 235
 DB 196 ACGITAVLVNASVSCP-----NHFEGLFRYKSIPEVDNQWEISAMFGQALGFIWMVKN 248
 QY 236 SNGCVLVHCLAGISRSATIAIATYMKRMDSLDEARYFYVEKPTTSPNNFGLDYR 295
 DB 249 SSGRVLVHCQAGISRSATICLAVLMQSRVRLDEAFVQKRGVISPNSFWGQLQFE 308
 QY 296 KKI 298
 DB 309 TVV 311
 RESULT 15
 DUS1_HUMAN STANDARD; PRT; 367 AA.
 AC P28562;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
 DE MAP kinase phosphatase-1 (MKP-1) (Protein-tyrosine phosphatase
 DE CL100) (Dual specificity protein phosphatase hvh1).
 GN DUSP1 OR PTPN10 OR MKP1 OR CL100 OR VH1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Forebrain;
 RX MEDLINE=93024952; PubMed=1406996;
 RA Keyes S.M., Emalie E.A.;
 RT "Oxidative stress and heat shock induce a human gene encoding a
 RT protein-tyrosine phosphatase.";
 RL Nature 359:644-647(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
 RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 RA Fahy J., Hulton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Botterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
 CC kinase ERK2 on both Thr-183 and Tyr-185.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- INDUCTION: By oxidative stress and heat shock.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL: X68277; CAA48338.1; -.
 CC EMBL: BC022463; AAH22463.1; -.
 DR PIR: S29090; S29090.
 DR HSP: Q16828; IMKP.
 DR GeneW: HGNC:3064; DUSP1.
 DR MIM: 600714; -.
 DR GO: GO:0004726; P:non-membrane spanning protein tyrosine phos. .; TAS.
 DR GO: GO:0006979; P:response to oxidative stress; TAS.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00195; DSPC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PSS0206; RHODANES_3; 1.
 DR PROSITE: PSS0383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PSS0056; TYR_PHOSPHATASE_2; 1.

DR PROSITE: PSS0054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolase; Cell cycle-
 FT DOMAIN 20 137
 FT DOMAIN 175 367
 FT ACT_SITE 258 258
 FT ACT_SITE 258 258
 FT ACT_SITE 258 258
 SO SEQUENCE 367 AA; 39297 MW; 11BD1D39A9FCD51F CRC64;
 Query Match 26.5%; Score 415; DB 1; Length 367;
 Best Local Similarity 33.9%; Pred. No. 3.6e-28;
 Matches 102; Conservative 47; Mismatches 134; Indels 18; Gaps 4;
 QY 15 LVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQODKYLITELIHSK 74
 DB 13 LRALDGERAAQCLLDRCSEFPAPNAGHAGSVNRSTIVRRRAKAMGLHIVPAE 72
 QY 75 HKVDIDCSQKVVVYDQSSQDVASLSQDCEFLVLLGRL--EKSPNVHLLAGFAEFSRCP 132
 DB 73 GRLLAGAVYAVVLLDERSAALDQAKKRDGTLLAAGALCREARAQVFLKGYEAFSASC 132
 QY 133 PGLCEK-----STLVPLCISQPLPVANT-----GPRILIPNLYLGGQRLVLR 177
 DB 133 PELCSKQSTPMGLSLPLSTSVDSASGSCSTPLDQGGVEILPFLYLSAVHASRK 192
 QY 178 ELIQNGIGVNLNASTYCKPDPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKSV 237
 DB 193 DMIDALGITLALINVSANCN--HFEHGYQKSLPVEDNHKADISSWNEAIDFIDSIKNG 251
 QY 238 GCVLVHCLAGISRSATIALAYIMKRMDSLDEAYRPFKEKRPITSPNPNLQGLDYKK 297
 DB 252 GRVFNVCQAGISRSATICLAYIMRTVRVLDFAEFVKQRRIISNFSFMGLQFESQ 311
 QY 298 I 298
 DB 312 V 312

Search completed: June 21, 2004, 13:21:07
 Job time : 6.55843 secs

This Page Blank (uspto)


```
QY 1 MAHEMIGTQIVTERLVALLESCTEKVLLIDSRPFVEYNTSHIIIEAININCSKIMKRRLQ 60
DB 1 MAHEMIGTQIVTERLVALLESCTEKVLLIDSRPFVEYNTSHIIIEAININCSKIMKRRLQ 60
QY 61 DKVITLTELQHSANRKHVDIDCSQKVVYVDSQDVASISSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVITLTELQHSANRKHVDIDCSQKVVYVDSQDVASISSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCGKSTLVPTCISQPCLPVANIPTIRLLPMLYLGQQRVNLKELI 180
DB 121 LAGGFAEFSRCFPGLCGKSTLVPTCISQPCLPVANIPTIRLLPMLYLGQQRVNLKELI 180
QY 181 QONGIGYVLNASNTCPKDPFIPESHFLRVVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLNASNTCPKDPFIPESHFLRVVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPENFLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPENFLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
```

RESULT 2

```
Q96N49 PRELIMINARY; PRT; 665 AA.
ID 096N49
AC 096N49
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ31411.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Taahiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato K., Nakamitsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Nagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima K., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Maeno Y., Nagai K.,
RA Isogai T.;
RT "MEDO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB051573; BAB71060.1; -.
DR HSSP; Q16828; IMKP.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0017017; P:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein; Hydrolyase.
SQ SEQUENCE 665 AA; 73058 MW; 1EABDFF08460DPF CRC64;
```

Query Match 98.8%; Score 1545; DB 4; Length 665;
Best Local Similarity 99.0%; Pred. No. 2.1e-137;
Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MAHEMIGTQIVTERLVALLESCTEKVLLIDSRPFVEYNTSHIIIEAININCSKIMKRRLQ 60
DB 1 MAHEMIGTQIVTERLVALLESCTEKVLLIDSRPFVEYNTSHIIIEAININCSKIMKRRLQ 60
QY 61 DKVITLTELQHSANRKHVDIDCSQKVVYVDSQDVASISSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVITLTELQHSANRKHVDIDCSQKVVYVDSQDVASISSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCGKSTLVPTCISQPCLPVANIPTIRLLPMLYLGQQRVNLKELI 180
DB 121 LAGGFAEFSRCFPGLCGKSTLVPTCISQPCLPVANIPTIRLLPMLYLGQQRVNLKELI 180
QY 181 QONGIGYVLNASNTCPKDPFIPESHFLRVVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLNASNTCPKDPFIPESHFLRVVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPENFLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPENFLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
```

RESULT 3

```
Q920R2 PRELIMINARY; PRT; 660 AA.
ID 0920R2
AC 0920R2
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE MAP kinase phosphatase-7.
GN DUSP16 OR 3830417M17RK OR MKP-7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Cerebellum;
RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.,
RT "MKP-7, a Novel Mitogen-activated Protein Kinase Phosphatase,
RT Functions as a Shuttle Protein."
RL J. Biol. Chem. 276:39002-39011 (2001).
DR EMBL; AB052157; BAB47240.1; -.
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0000188; F:inactivation of MAPK; IDA.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; kinase.
SQ SEQUENCE 660 AA; 72695 MW; DB609PCDAD4AA309 CRC64;
```

Query Match 97.0%; Score 1517; DB 11; Length 660;
Best Local Similarity 96.0%; Pred. No. 9.3e-135;
Matches 290; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 MAHEMIGTOIVTERLVALLSEGTKEYLLIDSRPVEYNTSHILEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTOIVTERLVALLSEGTKEYLLIDSRPVEYNTSHILEAININCSKLMKRRLQ 60
QY 61 DKVLITTELIQSHAKHAKVIDCSQKVVYVYDQSSODVASLSDDCFVTLGLKLEKSPNSVHL 120
DB 61 DKVLITTELIQSHAKHAKVIDCSQKVVYVYDQSSODVASLSDDCFVTLGLKLEKSPNSVHL 120
QY 121 LAGGFAEFSCFPGLCEGKSTLVPTCISQCLPVANIGPTRILPNIYLGCGQDVLNKL 180
DB 121 LAGGFAEFSCFPGLCEGKSTLVPTCISQCLPVANIGPTRILPNIYLGCGQDVLNKL 180
QY 181 QONGIGYVLANASTCPKPDPIPSHFLRPVNDSPCEKILPWLDSVDFTEKAKSNGCV 240
DB 181 QONGIGYVLANASTCPKPDPIPSHFLRPVNDSPCEKILPWLDSVDFTEKAKSNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTISPNFNLGQLDYEEKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTISPNFNLGQLDYEEKIKN 300
QY 301 QT 302
DB 301 QT 302
```

```
RESULT 4
Q99MG6 PRELIMINARY; PRT; 677 AA.
ID Q99MG6;
AC Q99MG6;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Map kinase phosphatase-M A1 isoform.
GN DUSP16 OR 3830417M17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.,
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages."
RL Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL: AF345951; AAK35052.1; -.
DR HSSP: Q16828; IMKP.
DR MCD; MGI:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHRASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PSS0206; RHODANSE_3; 1.
DR PROSITE; PSS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PSS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; Kinase.
SQ SEQUENCE 677 AA; 74550 MW; 8B6D5B7096CEC2FC CRC64;
```

Query Match 96.8%; Score 1514; DB 11; Length 677;
Best Local Similarity 95.7%; Pred. No. 1,9e-134;
Matches 289; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 MAHEMIGTOIVTERLVALLSEGTKEYLLIDSRPVEYNTSHILEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTOIVTERLVALLSEGTKEYLLIDSRPVEYNTSHILEAININCSKLMKRRLQ 60
QY 61 DKVLITTELIQSHAKHAKVIDCSQKVVYVYDQSSODVASLSDDCFVTLGLKLEKSPNSVHL 120
DB 61 DKVLITTELIQSHAKHAKVIDCSQKVVYVYDQSSODVASLSDDCFVTLGLKLEKSPNSVHL 120
QY 121 LAGGFAEFSCFPGLCEGKSTLVPTCISQCLPVANIGPTRILPNIYLGCGQDVLNKL 180
DB 121 LAGGFAEFSCFPGLCEGKSTLVPTCISQCLPVANIGPTRILPNIYLGCGQDVLNKL 180
QY 181 QONGIGYVLANASTCPKPDPIPSHFLRPVNDSPCEKILPWLDSVDFTEKAKSNGCV 240
DB 181 QONGIGYVLANASTCPKPDPIPSHFLRPVNDSPCEKILPWLDSVDFTEKAKSNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTISPNFNLGQLDYEEKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTISPNFNLGQLDYEEKIKN 300
QY 301 QT 302
DB 301 QT 302
```

```
RESULT 5
Q9AG16 PRELIMINARY; PRT; 355 AA.
ID Q9AG16;
AC Q9AG16;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Map kinase phosphatase-M B1 isoform.
GN DUSP16 OR 3830417M17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.,
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages."
RL Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL: AF345953; AAK35054.1; -.
DR HSSP: Q16828; IMKP.
DR MCD; MGI:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHRASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PSS0206; RHODANSE_3; 1.
DR PROSITE; PSS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PSS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; Kinase.
SQ SEQUENCE 355 AA; 39502 MW; D52C29AE215CA285 CRC64;
```

Query Match 87.9%; Score 1375.5; DB 11; Length 355;
Best Local Similarity 94.6%; Pred. No. 1e-121;
Matches 265; Conservative 8; Mismatches 4; Indels 3; Gaps 1;

QY 1 MAHEMIGTOIVTERLVALLESSTGTEKVLIDSRPFVEYNTSHILEANINCSKLMRRLOQ 60
DB 1 MAHEMIGTOIVTERLVALLESSTGTEKVLIDSRPFVEYNTSHILEANINCSKLMRRLOQ 60
QY 61 DKVLTIELIIOHSAKHVDIDCSQKVVYVYDSSQDVASLSDDCFLLVLLGKLEKRSFNSVHL 120
DB 61 DKVLTIELIIOHSAKHVDIDCSQKVVYVYDSSQDVASLSDDCFLLVLLGKLEKRSFNSVHL 120
QY 121 LAGGFABSRCPFGLCCEGKSTLVPTCISQPCLPVANIGPTRLIPMLYLGCQDVANKELEI 180
DB 121 LAGGFABSRCPFGLCCEGKSTLVPTCISQPCLPVANIGPTRLIPMLYLGCQDVANKELEI 180
QY 181 QONGIGVYVNASYTCPKDPFIPESHFLRPVNDSCCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGVYVNASYTCPKDPFIPESHFLRPVNDSCCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTPTSPNPNFLG 289
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTPTSPNPNFLG 289

RESULT 6

Q8N5T1 PRELIMINARY; PRT; 616 AA.
ID Q8N5T1
AC Q8N5T1
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DS Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_Taxid=9606;
[1]
SEQUENCE FROM N.A.
RC TISSUE=Brain, and lung;
RA Strauberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031643; AAH31643.1; -
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; F:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR000343; MAPK phosph.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS0054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 616 AA; 67636 MW; 2CB0B14482F2AD72 CRC64;

Query Match 83.9%; Score 1312; DB 4; Length 616;
Best Local Similarity 99.2%; Pred. No. 2.2e-115;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 50 CSKLMRRLOQDKVLTIELIIOHSAKHVDIDCSQKVVYVYDSSQDVASLSDDCFLLVLLG 109
DB 1 CSKLMRRLOQDKVLTIELIIOHSAKHVDIDCSQKVVYVYDSSQDVASLSDDCFLLVLLG 109
QY 110 KLEKRSFNSVHLLAGGFABSRCPFGLCCEGKSTLVPTCISQPCLPVANIGPTRLIPMLYL 169
DB 110 KLEKRSFNSVHLLAGGFABSRCPFGLCCEGKSTLVPTCISQPCLPVANIGPTRLIPMLYL 169
QY 61 KLEKRSFNSVHLLAGGFABSRCPFGLCCEGKSTLVPTCISQPCLPVANIGPTRLIPMLYL 120
DB 61 KLEKRSFNSVHLLAGGFABSRCPFGLCCEGKSTLVPTCISQPCLPVANIGPTRLIPMLYL 120
QY 170 QORDVANKELIQONGIGVYVNASYTCPKDPFIPESHFLRPVNDSCCEKILPMLDKSVDF 229
DB 170 QORDVANKELIQONGIGVYVNASYTCPKDPFIPESHFLRPVNDSCCEKILPMLDKSVDF 229
QY 121 QORDVANKELIQONGIGVYVNASYTCPKDPFIPESHFLRPVNDSCCEKILPMLDKSVDF 180
DB 121 QORDVANKELIQONGIGVYVNASYTCPKDPFIPESHFLRPVNDSCCEKILPMLDKSVDF 180
QY 230 IEKAKASNGCVLVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTPTSPNPNFLG 289
DB 230 IEKAKASNGCVLVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTPTSPNPNFLG 289

DB 181 IEKAKASNGCVLVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTPTSPNPNFLG 240
QY 290 QLDIEKKIKNOT 302
DB 241 QLDIEKKIKNOT 253

RESULT 7

Q99NG5 PRELIMINARY; PRT; 622 AA.
ID Q99NG5
AC Q99NG5
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Map kinase phosphatase-M A2 isoform.
GN DUSP16 OR 3830417M17R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
[1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Matsunuchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.,
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages.";
RL Mol. Cell. Biol. 20:6999-7009 (2001).
DR EMBL; AF345952; AAK35053.1; -
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPT.
DR GO; GO:0000188; F:inactivation of MAPK; IDA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR000343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00206; RHODANES 3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS0054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydroxylase, Kinase.
SQ SEQUENCE 622 AA; 68672 MW; 6C0CA4BEB9099B98 CRC64;

Query Match 75.0%; Score 1172.5; DB 11; Length 622;
Best Local Similarity 77.5%; Pred. No. 3.5e-102;
Matches 234; Conservative 8; Mismatches 5; Indels 55; Gaps 1;

QY 1 MAHEMIGTOIVTERLVALLESSTGTEKVLIDSRPFVEYNTSHILEANINCSKLMRRLOQ 60
DB 1 MAHEMIGTOIVTERLVALLESSTGTEKVLIDSRPFVEYNTSHILEANINCSKLMRRLOQ 60
QY 61 DKVLTIELIIOHSAKHVDIDCSQKVVYVYDSSQDVASLSDDCFLLVLLGKLEKRSFNSVHL 120
DB 61 DKVLTIELIIOHSAKHVDIDCSQKVVYVYDSSQDVASLSDDCFLLVLLGKLEKRSFNSVHL 120
QY 121 LAGGFABSRCPFGLCCEGKSTLVPTCISQPCLPVANIGPTRLIPMLYLGCQDVANKELEI 180
DB 121 LAGGFABSRCPFGLCCEGKSTLVPTCISQPCLPVANIGPTRLIPMLYLGCQDVANKELEI 180
QY 181 QONGIGVYVNASYTCPKDPFIPESHFLRPVNDSCCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGVYVNASYTCPKDPFIPESHFLRPVNDSCCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTPTSPNPNFLGQLDIEKKIKNOT 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTPTSPNPNFLGQLDIEKKIKNOT 300

Db 186 LHHLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTISPNFNFQMLMDEKTIIN 245
Qy 301 QT 302
Db 246 QT 247

RESULT 8
ID Q9AG15 PRELIMINARY; PRT; 300 AA.
AC Q9AG15;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Map kinase phosphatase-M B2 isoform.
GN DUSP16 OR 383041M1/RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Matsunuchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.;
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages";
RL Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL; AF345954; AAK35055.1; -.
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR00387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00450; RHOD; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolyase; Kinase.
SQ SEQUENCE 300 AA; 33624 MW; 4C61846ACDF0F456 CRC64;

Query Match 66.1%; Score 1034; DB 11; Length 300;
Best Local Similarity 75.0%; Pred. No. 1.8e-89;
Matches 210; Conservative 8; Mismatches 4; Indels 58; Gaps 2;

Qy 1 MAHEMIGTQV-TERLVALLSSTGTFKULLIDSRPFVEVNTSHILEAININCSKLMKRLQ 60
Db 1 MAHEMIGTQVTERLVALLSSTGTFKULLIDSRPFVEVNTSHILEAININCSKLMKRLQ 60
Qy 61 DKVLITELIHSASAKKVIDCSQKVVVYDDOSSQDVASLSSDCFLTVLLGKLEKSFNSVH 120
Db 61 DKVLITELIHSASAKKVIDCSQKVVVYDDOSSQDVASLSSDCFLTVLLGKLEKSFNSVH 120
Qy 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIPTRIIPNLVYLGQRDVANKELI 180
Db 121 LA-----DLM 125
Qy 181 QONGIGVVLNASTCPKPDPIRESHPLRVPNDSFCEKILPMLDKSVDFLEKASNGCV 240
Db 126 QONGIGVVLNASTCPKPDPIRESHPLRVPNDSFCEKILPMLDKSVDFLEKASNGCV 185
Qy 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPT 280
Db 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPT 280

Db 186 LHHLAGISRSATIAIAYIMKMDMSLDEAYR---RQKPT 222

RESULT 9
ID Q86SS8 PRELIMINARY; PRT; 625 AA.
AC Q86SS8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to dual specificity phosphatase 8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Astrocytoma;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045110; AA045110.1; -.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR002965; P rich-extension.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR PRINTS; PRO1217; PRICEXTENS.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 625 AA; 65826 MW; C7C608407B724FPC CRC64;

Query Match 63.2%; Score 989; DB 4; Length 625;
Best Local Similarity 63.1%; Pred. No. 8.3e-85;
Matches 190; Conservative 55; Mismatches 54; Indels 2; Gaps 2;

Qy 1 MAHEMIGTQV-TERLVALLSSTGTFKULLIDSRPFVEVNTSHILEAININCSKLMKRLQ 59
Db 1 MAGDRPRKVMQKXKLLASLGGPGLVYDSSFEVYNSWHVLSVNICCSKLVKRLQ 60
Qy 60 QDVLITELIHSASAKKVIDCSQKVVVYDDOSSQDVASLSSDCFLTVLLGKLEKSFNSVH 119
Db 61 QGKVTIELIHSASAKKVIDCSQKVVVYDDOSSQDVASLSSDCFLTVLLGKLEKSFNSVH 120
Qy 120 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIPTRIIPNLVYLGQRDVANKELI 178
Db 121 ILTGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIPTRIIPNLVYLGQRDVANKELI 180
Qy 179 LQONGIGVVLNASTCPKPDPIRESHPLRVPNDSFCEKILPMLDKSVDFLEKASNG 238
Db 181 LQONGIGVVLNASTCPKPDPIRESHPLRVPNDSFCEKILPMLDKSVDFLEKASNG 240
Qy 239 CVLVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTISPNFNFQMLMDEKTIIN 298
Db 241 QVHVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTISPNFNFQMLMDEKTIIN 300
Qy 299 K 299
Db 301 K 301

RESULT 10
Q7TS29 PRELIMINARY; PRT; 665 AA.
ID Q7TS29;
AC Q7TS29;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TReMBLrel. 25, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE Dusp8 protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., McEwan N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Schaefer C.F., Bhat N.K.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., McKernan K.J., Mullikin J.C.,
 RA Krzywinski M.I., Skalski U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC052705; AAHS2705.1; -
 SQ SEQUENCE 665 AA; 69021 MW; 9166E36A835249F CRC64;

Query Match 63.0%; Score 985; DB 11; Length 665;
 Best Local Similarity 62.5%; Pred. No. 2.2e-84;
 Matches 188; Conservative 55; Mismatches 56; Indels 2; Gaps 2;

QY 1 MAHEMGTQIV-TERLVALLSGTEKVLIDSRPFVEYNTSHILEAININSGKMKRRLOQ 59
 DB 1 MADGRLEPRKVMADKXIASLIRGPGPLVIDSRPFVEYNTSHILEAININSGKMKRRLOQ 60
 QY 60 QDVKLTTELIOHSAKHKVINDCSOKVYVYDSSQVVASISDCFTLVLLGLKLEKSFNSVH 119
 DB 61 QGKVTIAELIOPATRSQVATEPQDVVYDOSTRASVLAADSFLLSLKLDGCFDPSVA 120
 QY 120 LLAGFAEBSRCPPGLCEGKSTLVPT-CISQPCLPVANIPTRIIPNLVYLGCCQDVYLNKE 178
 DB 121 ILTGGFATPSSCFPGICEGKPATLPSMELSPGCLPVPSPVGLTRILPHLYLSQSDVYLNKD 180
 QY 179 LTIQNGIGVNLASTYCPKDPPIPSHSLRPVNDSEFEKLLPWLKDSVDFIEKAKASNG 238
 DB 181 LMTQNGISVNLANSKCPDPTICSRFRWRIPINDNYCEKLLPLDKSIFIDIKAKLSSC 240
 QY 239 CVLVHCLAGISRSATIALAYIMKMDMSIDEAYFVEKEKPTSPNNPFGOLLDVEKXI 298
 DB 241 QVIVHCLAGISRSATIALAYIMKMGMSSDAYFVDRKPSISPNENFLGQLEVERSL 300
 QY 299 K 299
 DB 301 K 301

RESULT 11
 Q8BZE4 PRELIMINARY; PRT; 206 AA.
 AC Q8BZE4;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE Truncated MAPK phosphatase 7 homolog.
 GN DEERD213E
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=urinary bladder;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL: AK035652; BAC29138.1; -
 DR MGD: MGI:1098836; DEERD213E.
 DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR001763; Rhodanese-like.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00450; RHOD, 1.
 DR PROSITE: PSS0206; RHODANESE 3; 1.
 SQ SEQUENCE 206 AA; 22566 MW; E7PDE5C01ADC2F73 CRC64;

Query Match 57.0%; Score 891; DB 11; Length 206;
 Best Local Similarity 97.7%; Pred. No. 3.8e-76;
 Matches 173; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININSGKMKRRLOQ 60
 DB 1 MAHEMGTQIVTESLVALLSGTEKVLIDSRPFVEYNTSHILEAININSGKMKRRLOQ 60
 QY 61 DKVLITELIOHSAKHKVINDCSOKVYVYDSSQVVASISDCFTLVLLGLKLEKSFNSVH 120
 DB 61 DKVLITELIOHSAKHKVINDCSOKVYVYDSSQVVASISDCFTLVLLGLKLEKSFNSVH 120
 QY 121 LLAGFAEBSRCPPGLCEGKSTLVPTCISQPCLPVANIPTRIIPNLVYLGCCQDVYLNKE 177
 DB 121 LAGGFAEBSRCPPGLCEGKSTLVPTCISQPCLPVANIPTRIIPNLVYLGCCQDVYLNKE 177

RESULT 12
 Q96Q52 PRELIMINARY; PRT; 143 AA.
 AC Q96Q52;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE Truncated MAPK phosphatase 7.
 GN MKP7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Montpetit A., Boly G., Simmet D.,
 RT "A detailed transcriptional map of the chromosome 12p12 tumor
 suppressor locus.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY038927; AAK69770.1; -
 DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00450; RHOD, 1.
 DR PROSITE: PSS0206; RHODANESE 3; 1.

DR PROSITE; PSS0206; RHODANSE 3; 1.
SQ SEQUENCE 143 AA; 16077 MW; 5213A213AA7E5974 CRC64;
Query Match 39.0%; Score 610; DB 4; Length 143;
Best Local Similarity 98.4%; Pred. No. 9.5e-50;
Matches 124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAHEMGTQVTEKRLVALLSGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
DB 1 MAHEMGTQVTEKRLVALLSGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
QY 61 DKYLITELIQSHAKHKVDIDCSQKVVVYDQSSQDVASLSDCFLTVLLGLKESFNSVHL 120
DB 61 DKYLITELIQSHAKHKVDIDCSQKVVVYDQSSQDVASLSDCFLTVLLGLKESFNSVHL 120
QY 121 LAGGFA 126
DB 121 LAGADA 126
RESULT 13
Q8ST19 PRELIMINARY; PRT; 657 AA.
ID Q8ST19
AC Q8ST19;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F08B1.1a.
GN F08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RX Chisase S.;
RA "The sequence of C. elegans cosmid F08B1.";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Waterston R.;
RT "Direct Submision.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR HSSP: Q16828; IMKP.
DR WormPep; F08B1.1a; CE27918.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:Protein amino acid dephosphorylation; IEA.
DR InterPro; IPRO00340; DS phosphatase.
DR InterPro; IPRO00343; MAPK phosph.
DR InterPro; IPRO01763; Rhodanese-like.
DR InterPro; IPRO00387; TYR phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PSS0206; RHODANSE 3; 1.
DR PROSITE; PSS0383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein; Hydrolase.
SQ SEQUENCE 657 AA; 71002 MW; A0D9153DBE6326B43 CRC64;

Query Match 36.0%; Score 563.5; DB 5; Length 657;
Best Local Similarity 42.0%; Pred. No. 1.7e-44;
Matches 131; Conservative 61; Mismatches 87; Indels 33; Gaps 10;
QY 10 IYTERVALLSGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQODRV---LI 65
DB 9 ISTGIALIIRPDPITLVVDCRGFEYNEISHVHSHMNAFSLIRRLRFENLIDNCLI 68
QY 66 TELIQSHA-----KHKVDIDCSQKVVVYDQ-----SSQDVASL-----SSDCLTVLLGK 110
DB 69 HQLMSCSSGCTKMDKDLT-----VLVAEDKPRGNKRIASCNAPESAKIMRVLRER 122
QY 111 LEKS--FNSVHLLAGFAFARFCRPGLCGKS--TLVPCISQPC--PYANIGPRILP 164
DB 123 LEDTDFRSVMVLEGGFKQPAQOYPOLCESSSEGMTPLPSLSQPCLSQPGD-GITLITP 181
QY 165 NLVLCGQRDVLNKEILIQNGIGVLNASYTCPPKDFIPE-SHFLRPVNDSPCEKILPWL 223
DB 182 NITLGSQIDSLDEMTLDALDISVINLSMTCPKSCVICEDKKNFMRIPVNDYQKLSPIY 241
QY 224 DKSVDPIEKAKASNGCVLVHCLAGISRSATIAIYIMKMDSLDAVYRVEKEKPTISP 283
DB 242 PMAYEFLEKCRPRAGKCLHCLAGISRSPTLAISYIMRYMKGSDDAYRYKRRPSISP 301
QY 284 NRPVPLGQLIDYE 295
DB 302 NPFNMQLLLEYE 313
RESULT 14
Q7SZF3 PRELIMINARY; PRT; 367 AA.
ID Q7SZF3
AC Q7SZF3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shennan C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;
RA Brownstein M.J.; Udén T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA Raha S.S.; Loggellano N.A.; Peters G.J.; Abramson R.D.; Mullany S.J.;
RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
RA Villalón D.K.; Muny D.M.; Sodeguren E.J.; Lu X.; Gibbs R.A.;
RA Fahy J.; Helton E.; Ketteman M.; Madan A.; Rodriguez S.; Sanchez A.;
RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Buttefield Y.S.;
RA Krzywinski M.I.; Skalska U.; Smalins D.E.; Schmech A.; Schein J.E.;
RA Jones S.J.; Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052477; AAH52477.1; -.

KW Hypothetical protein.
SQ SEQUENCE 367 AA; 40538 MW; F89997A415DAD6AC CRC64;
Query Match 30.2%; Score 472.5; DB 13; Length 367;
Best Local Similarity 34.8%; Pred. No. 3.3e-36;
Matches 106; Conservative 66; Mismatches 110; Indels 19; Gaps 6;
QY 15 LVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKMKRR---LOODKVLITELIQ 70
DB 15 LKRLMKDDGAKCLLDORSLFASAGHLGAVINRCNTIVRRRAKSVSLDQILSD--- 71
QY 71 HSAKHKVIDCSQKVVVYDQSSQDVASLSDCFLTVLGLK-EKSFNS-VHLLAGGFAEF 128
DB 72 DEARSRKSGLSYAVILYDERSSDITVTMKDSTITVHVALCRDFTSTEVYLLKGGYDRF 131
QY 129 SRCEPGLC-----EGKSTLVPTCISQPCLPVANIPTRIILPVLVYCCQDVIANEL 179
DB 132 STQYPRYCKLTKRTLSVSSQSSMESCSCATPQHDGGPVELPLFLGSLALHASKDM 191
QY 180 IQONGIGVYLNASVYCPKPDFIPESHFLRPVNDSPCEKILPWLKSVDFIEKASNGC 239
DB 192 LDMGISALINVSNCPPN-HFEGDYQKCIPEVDNKHEDISSMFLBAIEFIDSVDSNGR 250
QY 240 VLVHCLAGISRSATIIAYIMKRMDSLEAYRPFYKERRPTISPNPFLGQLLDYEKKI 298
DB 251 VLVHCGAGISRSATICTLAYIMKKRVRLAEAFEFYKORRSITISPNPFLGQLDPESCV 309
RESULT 15
ID 08R3L3 PRELIMINARY; PRT; 483 AA.
AC 08R3L3;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Dual specificity phosphatase 10.
GN DUSP10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN 11
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RP [2]
RC STRAIN=C57BL/6J, and MOD;
RC TISSUE=Pancreas, Thymus, and Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RT EMBL; BC025066; AAH25066.1;
DR EMBL; AK035293; BAC29019.1;
DR EMBL; AK050528; BAC34308.1;
DR EMBL; AK088024; BAC40102.1;
DR EMBL; AK088186; BAC40196.1;
DR EMBL; AK088357; BAC40300.1;
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1927070; Dusp10.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR003440; DS phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.

DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
SQ SEQUENCE 483 AA; 52532 MW; 7797A1877D986AAB CRC64;
Query Match 30.0%; Score 469; DB 11; Length 483;
Best Local Similarity 35.6%; Pred. No. 1e-35;
Matches 106; Conservative 67; Mismatches 95; Indels 30; Gaps 7;
QY 27 ILIDSRPFVEYNTSHLEAININCS-KLMKRRIQODKVLITELIQ-HSAKHKVIDCSQK 84
DB 174 VTIIDCRPFMEYNKSHIQGAVHINCAKISRRRLQGGKITVLDISCREGDSFKRLFSKE 233
QY 85 VVYVDQSSQDVASLSDCFLTVLGLKLEKSFNSVHLLAGGFAEFRCFPGLCEGK----- 139
DB 234 IIVYDENTNPSRVTSPQPLHIVLESIKRGEPLVYKGLSFKQNHGRLCDNSIQLOE 293
QY 140 -----STLVPTCISQPCLP-VANIGPTRLIIPNLVYCCQDVIANKELIQONGIGY 187
DB 294 CREYGGGASNAASSMLPQ--SVPTTPIENAEILPILPFLFLGNEQDAQDLDTMQRLNIGY 351
QY 188 VLNAS-----YTCPKPDFIPESHFLRPVNDSPCEKILPWLKSVDFIEKASNCVIV 242
DB 352 VINVTTHLPVLYHKEGLF---NYKRLPATDSKKNLRQYFBEAFEFBEAHQCGKGLIT 407
QY 243 HCLAGISRSATIIAYIMKRMDSLEAYRPFYKERRPTISPNPFLGQLLDYEKKIKN 300
DB 408 HCGAGVSRSAITIVAYIMGTRMTMDAYKFPYKGRKPTISPNLNTMFGQLLEFEEDLNN 465

Search completed: June 21, 2004, 13:22:36
Job time : 24.7353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:10:27 ; Search time 33.7291 Seconds

(without alignments)
2529.847 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302

Perfect score: 1564
Sequence: 1 MAHEMIGTQVTERIVALL.....PNFPIGQLDYEKIKKINQI 302

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_29Jan04:*

1: geneseqp1980s:*\n2: geneseqp1980s:*\n3: geneseqp2000s:*\n4: geneseqp2000s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1564	100.0	665	5	ABR52381 Protein r
2	1564	100.0	665	5	ABR52407
3	1552	99.2	665	4	AAE04834
4	1552	99.2	665	4	AAU09016
5	1552	99.2	665	5	AAU79156
6	1552	99.2	665	5	AAU09946
7	1552	99.2	665	5	AAU75789
8	1552	99.2	665	5	ABR97946
9	1552	99.2	665	5	AAU79929
10	1552	99.2	665	5	ABR97291
11	1545	98.8	665	6	ADAS4744
12	1544	98.7	665	5	ABR52352
13	1544	98.7	665	5	ABR52352
14	1544	98.7	665	5	ABR52352
15	1544	98.7	665	5	ABR52352
16	1542	98.6	665	5	AAU79162
17	1526.5	97.6	664	5	ABR52424
18	1492	95.4	660	5	ABR52385
19	988	63.2	625	5	ABR52382
20	988	63.2	625	5	ABR52350
21	988	63.0	625	6	AAW29150
22	985	63.0	663	5	ABR52351
23	985	63.0	663	5	ABR52351
24	985	63.0	663	5	ABR52351
25	742.5	47.5	375	4	ABG07902

26	741	47.4	140	5	ABR52404	ABR52404 Peptide r
27	716	45.8	140	5	ABR52405	ABR52405 Peptide r
28	715.5	45.7	579	7	ADP08458	ADP08458 Novel pro
29	654.5	41.8	170	4	AAE06780	AAE06780 Human MAP
30	654.5	41.8	170	4	AAE06780	AAE06780 Human MAP
31	646	41.3	517	5	AAU79159	AAU79159 Human dua
32	476	30.4	482	4	AAE73225	AAE73225 Human pho
33	469	30.0	444	5	AAO20515	AAO20515 Protein o
34	469	30.0	482	3	AAE18655	AAE18655 A human r
35	469	30.0	482	3	AAE29641	AAE29641 Human dua
36	469	30.0	482	5	ABR52384	ABR52384 Protein r
37	469	30.0	482	5	ABR52384	ABR52384 Protein r
38	467.5	29.9	381	5	ABR52383	ABR52383 Breast ca
39	467.5	29.9	381	6	ABP96803	ABP96803 Human COP
40	462.5	29.6	381	7	ADP48300	ADP48300 Mouse MKP
41	461.5	29.5	381	7	ADP48300	ADP48300 Rat Prote
42	461.5	29.5	381	7	ADP48300	ADP48300 Rat Prote
43	440.5	28.2	419	6	ABP55026	ABP55026 Human dua
44	440.5	28.2	419	6	ABP55026	ABP55026 Human dua
45	426	27.2	394	4	AAE67627	AAE67627 Amino aci

ALIGNMENTS

RESULT 1
ABR52381
ID ABR52381 standard; protein; 665 AA.

AC ABR52381;

DT 19-JUN-2003 (first entry)

DE Protein relating to the invention SEQ ID NO: 109.

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;

CC the invention has antiproliferative, hepatotropic, nephrotropic, CC
CC antiarthritic, antiproliferative, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX
XX
SQ Sequence 665 AA;

Query Match 100.0%; Score 1564; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 3.4e-164;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTRRLVALLSGTEKVLIDSRPFVEYNNTSHLEAININCSKLMKRLQ 60
DB 1 MAHEMIGTQVTRRLVALLSGTEKVLIDSRPFVEYNNTSHLEAININCSKLMKRLQ 60
QY 61 DKVLTITELIQHSARKHVIDCSQKVVYDQSSQDVASLSDCFLTVLLGKLEKSPNSVHL 120
DB 61 DKVLTITELIQHSARKHVIDCSQKVVYDQSSQDVASLSDCFLTVLLGKLEKSPNSVHL 120
QY 121 LAGFAFERSRCPGCEGKSTLVPTCISQPCLPVANIPTRLIPMLYGCORVYLNKELI 180
DB 121 LAGFAFERSRCPGCEGKSTLVPTCISQPCLPVANIPTRLIPMLYGCORVYLNKELI 180
QY 181 QONGIGYVLANASYTCRPFIPESHFLRPVNDSPCEKILPMLDSDVPIEKAKANGCV 240
DB 181 QONGIGYVLANASYTCRPFIPESHFLRPVNDSPCEKILPMLDSDVPIEKAKANGCV 240
QY 241 LVHCLAGISRATIAIAYIMKMDSLDBAYRFVKEKPTTSPNPNFLGQLLDYKIKIN 300
DB 241 LVHCLAGISRATIAIAYIMKMDSLDBAYRFVKEKPTTSPNPNFLGQLLDYKIKIN 300
QY 301 QT 302
DB 301 QT 302

RESULT 2
ABR52407
ID ABR52407 standard; protein; 665 AA.

XX ABR52407;

DT 19-JUN-2003 (first entry)

XX Protein relating to the invention SEQ ID NO: 148.

XX antiproliferative; hepatotropic; nephrotropic; antiarthritic;

XX antiproliferative; cardiant; cytostatic; gene therapy; liver disease;

XX immunological disorder; arthritis; psoriasis; congenital heart defect;

XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

XX Homo sapiens.

XX WO200257460-A2.

XX 25-JUL-2002.

XX 20-DEC-2001; 2001MO-US050459.

XX 20-DEC-2000; 2000US-0256868P.

XX 30-MAR-2001; 2001US-0280186P.

XX 01-MAY-2001; 2001US-0287735P.

XX 05-JUN-2001; 2001US-0295848P.

XX 25-JUN-2001; 2001US-0300465P.

PA (BRM) BRISTOL-MYERS SQUIBB CO.
XX Jackson DG, Feder J, Nelson T, Mintler G, Ramanathan C, Lee L,
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
PI Krystek S, Mcatee P, Suchard S, Bans D,
XX WPI, 2002-599721/64.
DR N-PSDB; ACC60572.

PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.

PS Disclosure; Fig 19; 801pp; English.

CC The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40

CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic, CC
CC antiarthritic, antiproliferative, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX

SQ Sequence 665 AA;

Query Match 100.0%; Score 1564; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 3.4e-164;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTRRLVALLSGTEKVLIDSRPFVEYNNTSHLEAININCSKLMKRLQ 60
DB 1 MAHEMIGTQVTRRLVALLSGTEKVLIDSRPFVEYNNTSHLEAININCSKLMKRLQ 60
QY 61 DKVLTITELIQHSARKHVIDCSQKVVYDQSSQDVASLSDCFLTVLLGKLEKSPNSVHL 120
DB 61 DKVLTITELIQHSARKHVIDCSQKVVYDQSSQDVASLSDCFLTVLLGKLEKSPNSVHL 120
QY 121 LAGFAFERSRCPGCEGKSTLVPTCISQPCLPVANIPTRLIPMLYGCORVYLNKELI 180
DB 121 LAGFAFERSRCPGCEGKSTLVPTCISQPCLPVANIPTRLIPMLYGCORVYLNKELI 180
QY 181 QONGIGYVLANASYTCRPFIPESHFLRPVNDSPCEKILPMLDSDVPIEKAKANGCV 240
DB 181 QONGIGYVLANASYTCRPFIPESHFLRPVNDSPCEKILPMLDSDVPIEKAKANGCV 240
QY 241 LVHCLAGISRATIAIAYIMKMDSLDBAYRFVKEKPTTSPNPNFLGQLLDYKIKIN 300
DB 241 LVHCLAGISRATIAIAYIMKMDSLDBAYRFVKEKPTTSPNPNFLGQLLDYKIKIN 300
QY 301 QT 302
DB 301 QT 302

RESULT 3
AAE04834
ID AAE04834 standard; protein; 665 AA.

XX AAE04834;

DT 10-SEP-2001 (first entry)

XX Human GSP002 phosphatase polypeptide.

XX Human; GSP002 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; ocular disease; organ transplant rejection;

KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
 KW metabolic disorder; hematopoietic cancer; mood disorder; cardiac;
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
 KW attention disorder; cognition disorder; psychotic disorder; cytosolic;
 KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;
 KW neuroprotective; antibacterial; vulnary; tranquilliser; anasthetic;
 KW hypotensive; immunosuppressive; antipariatic; analgesic; hypertensive;
 KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
 KW MKP; migraine; chromosome 12p11.1-p12.1.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 PH Domain 1..173
 FT /label= Catalytic_domain
 FT Domain 158..297
 FT /label= Phosphatase_domain
 XX
 XX MO200146394-A2.
 XX
 PD 28-JUN-2001.
 XX
 XX 21-DEC-2000; 2000WC-US034736.
 XX
 XX 21-DEC-1999; 99US-0173255P.
 XX 28-DEC-1999; 99US-0175766P.
 PR 25-JAN-2000; 2000US-0178078P.
 PR 31-JAN-2000; 2000US-0179301P.
 XX
 XX (SUGB-) SUGEN INC.
 XX
 PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;
 PI Hill RJ, Flanagan P;
 DR MPI; 2001-418058/44.
 DR N-PSDB; AAD09492.
 XX
 PT Novel phosphatase polypeptide useful for treating cancers, immune-related
 PT diseases and disorders, cardiovascular disease, brain or neuronal-
 PT associated diseases and metabolic disorders.
 XX
 PS Claim 7; Fig 2; 186pp; English.
 XX
 CC The present invention relates to phosphatase polypeptides, nucleotide
 CC sequences encoding them, as well as various products and methods useful
 CC for the diagnosis and treatment of various phosphatase-related diseases
 CC and conditions. Substance that modulates the activity of phosphatase
 CC polypeptide is used to treat immune-related diseases and disorders,
 CC cardiovascular disease, brain or neuronal-associated diseases and
 CC hematopoietic disorders, including cancers of tissues, cancers of
 CC hematopoietic origin, diseases of central and peripheral nervous system,
 CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
 CC lateral sclerosis, viral infections, infections caused by prions,
 CC bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual
 CC dysfunction, mood disorders, attention disorders, neurological disorders,
 CC hypertension, psychotic disorders, neurologic disorders,
 CC dyskinesias and organ transplant rejection. The present amino acid
 CC sequence is human SCP002 phosphatase polypeptide. This sequence is
 CC classified as dual specificity phosphatase (DSP) and MAP kinase
 CC phosphatase (MKP). SCP002 gene maps to chromosomal position 12p11.1-p12.1
 XX
 SQ Sequence 665 AA;
 Query Match 99.2%; Score 1552; DB 4; Length 665;
 Best Local Similarity 99.3%; Pred. No. 7.4e-163;
 Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MAHEMIGTQIVTRLVALLBSGTEKVLIDSRPFVYNTSHILEAININCKMKRRLLQ 60
 DB 1 MAHEMIGTQIVTRLVALLBSGTEKVLIDSRPFVYNTSHILEAININCKMKRRLLQ 60
 QY 61 DKVLITELLOHSAKHKVNDIDSGKVVVYQSSQDVASLSDDCLVTYLTGLKLEKSFNSVHL 120

DB 61 DKVLITELLOHSAKHKVNDIDSGKVVVYQSSQDVASLSDDCLVTYLTGLKLEKSFNSVHL 120
 QY 121 LAGFAEPRRCPPGLCEGKSTLVPTGISQPCLPVANIQTTRILPNLYGCGQDVYLNKELI 180
 DB 121 LAGFAEPRRCPPGLCEGKSTLVPTGISQPCLPVANIQTTRILPNLYGCGQDVYLNKELI 180
 QY 181 QONGIGYVYNAAYTCPPKPFIPESHPLRVVNDSPCEKLLPMLDKSVDFTEKAKASNGCV 240
 DB 181 QONGIGYVYNAAYTCPPKPFIPESHPLRVVNDSPCEKLLPMLDKSVDFTEKAKASNGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTTISPENFLGQLLDYEKKIRN 300
 DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTTISPENFLGQLLDYEKKIRN 300
 QY 301 QT 302
 DB 301 QT 302
 RESULT 4
 ID AAU09016 standard; protein; 665 AA.
 XX
 AC AAU09016;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human dual specificity phosphatase 21117.
 XX
 KW Human; dual specificity phosphatase 21117; hepatotropic; cytosolic;
 KW hematopoietic disorder; autoimmune disorder; diabetes mellitus;
 KW rheumatoid arthritis; multiple sclerosis; Crohn's disease;
 KW liver disorder; erythroid associated disorder; haemolytic anaemia;
 KW cellular proliferative; differentiative disorder; leukaemia;
 KW acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer;
 KW immunogen.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 11..131
 FT /label= Rhodanese-like domain
 FT Domain 158..297
 FT /label= Catalytic domain
 FT /note= "Dual specificity phosphatase catalytic domain"
 FT Active-site 242..254
 FT /label= Tyrosine_specific_protein_phosphatase_active_site
 PN MO200173059-A2.
 XX
 PD 04-OCT-2001.
 XX
 PD 23-MAR-2001; 2001WC-US009477.
 XX
 PF 24-MAR-2000; 2000US-0191858P.
 XX
 PR (MILL-) MILLENNIUM PHARM INC.
 XX
 PA Meyers RA;
 XX
 PI MPI; 2001-611635/70.
 XX
 DR N-PSDB; AAS14639.
 DR
 XX
 PT New human dual specificity polypeptides and nucleic acids for diagnosis
 PT of disease and treatment of e.g. liver disorders.
 XX
 PS Claim 9; Fig 1; 143pp; English.
 XX
 CC The invention relates to two novel human dual specificity phosphatases
 CC designated 21117 and 38692, the nucleic acids encoding them (including
 CC fragments, allelic variants, their complements or nucleic acids that
 CC hybridise to them) and antibodies raised against the proteins. The

CC antibody is useful for detecting the presence of the polypeptide, and the
CC nucleic acid fragments are useful for detecting the presence of the
CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
CC antisense sequences) are useful for modulating the activity or expression
CC of the polypeptide/nucleic acid, useful for the treatment of e.g.
CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
CC listed in the specification) liver disorders, erythroid associated
CC disorders (e.g. haemolytic anaemia) cellular proliferative or
CC differentiative disorders, leukemias (e.g. acute myeloid leukaemia),
CC carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
CC 38692 are also useful for modulating the proliferation, survival,
CC migration or differentiation of a 38692 or 21117-expressing cell. The
CC polypeptide and nucleic acids are useful for identifying modulating
CC agents. The present sequence represents the dual specificity phosphatase
CC 21117

XX Sequence 665 AA;

Query Match 99.2%; Score 1552; DB 4; Length 665;
Best Local Similarity 99.3%; Pred. No. 7,4e-163; Indels 0; Gaps 0;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MAHEMIGTQIVTERLVALLESGTEKVLIDSRPFVEYNTSHLEAININCSKMKRRLOQ 60
DB 1 MAHEMIGTQIVTERLVALLESGTEKVLIDSRPFVEYNTSHLEAININCSKMKRRLOQ 60
QY 61 DKVLTTELIOHSAKHKVDIDCSQKVVVYDQSSQDVASISDPCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLTTELIOHSAKHKVDIDCSQKVVVYDQSSQDVASISDPCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTIILPNLYLGCQRDVLANKELI 180
DB 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTIILPNLYLGCQRDVLANKELI 180
QY 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTIILPNLYLGCQRDVLANKELI 180
DB 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTIILPNLYLGCQRDVLANKELI 180
QY 181 QONGIGYVLNASYTCRKPDPFIPESHFLRPVNDSCFCEKILPWLKSDVFIEKAKASNGCV 240
DB 181 QONGIGYVLNASYTCRKPDPFIPESHFLRPVNDSCFCEKILPWLKSDVFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
```

RESULT 5
AAU79156
AAU79156 standard; protein; 665 AA.

XX 02-JUL-2002 (first entry)

DE Human dual-specificity phosphatase-3 (DSP-16) protein.

XX Human, dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; chromosome 12p.

XX Homo sapiens.

OS Key Location/Qualifiers
FH Domain 242..251
FT /label=Active_site_domain

XX MO200226997-A2.

PD 04-APR-2002.
XX
XX 25-SEP-2001; 2001WO-US030124.
PF
XX 26-SEP-2000; 2000US-0235487P.
PR
XX (CEPT)- CEPTYR INC.
PA

XX Luche RM, Wei B;

XX WPI; 2002-315802/35.
DR
XX N-PSDB; ABK4759C.

PT New DSP-16 polypeptide, useful for identifying modulators of its
FT activity, which can be used in the treatment of disorders such as
PT Duchenne muscular dystrophy, or cancer.

PS Claim 1; Fig 2; 87pp; English.

CC The present invention relates to a new polypeptide, DSP-16, having a 665
CC amino acid sequence, given in the specification, or a variant having at
CC least 50 % identical residues, which retains the ability to
CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
CC invention can be used for identifying agents which modulate DSP-16
CC activity for modulation of a proliferative response in a cell, survival
CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne
CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
CC cell proliferation, and cell cycle abnormalities. The present amino acid
CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
CC protein of the invention. This sequence is encoded by the human DSP-16
CC gene located on chromosome 12p

XX Sequence 665 AA;

Query Match 99.2%; Score 1552; DB 5; Length 665;
Best Local Similarity 99.3%; Pred. No. 7,4e-163; Indels 0; Gaps 0;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MAHEMIGTQIVTERLVALLESGTEKVLIDSRPFVEYNTSHLEAININCSKMKRRLOQ 60
DB 1 MAHEMIGTQIVTERLVALLESGTEKVLIDSRPFVEYNTSHLEAININCSKMKRRLOQ 60
QY 61 DKVLTTELIOHSAKHKVDIDCSQKVVVYDQSSQDVASISDPCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLTTELIOHSAKHKVDIDCSQKVVVYDQSSQDVASISDPCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTIILPNLYLGCQRDVLANKELI 180
DB 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTIILPNLYLGCQRDVLANKELI 180
QY 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTIILPNLYLGCQRDVLANKELI 180
DB 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTIILPNLYLGCQRDVLANKELI 180
QY 181 QONGIGYVLNASYTCRKPDPFIPESHFLRPVNDSCFCEKILPWLKSDVFIEKAKASNGCV 240
DB 181 QONGIGYVLNASYTCRKPDPFIPESHFLRPVNDSCFCEKILPWLKSDVFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
```

RESULT 6
AAU09946
AAU09946 standard; protein; 665 AA.

XX AAU09946;

DT 18-JUN-2002 (first entry)
XX Protein sequence of human (dual specificity phosphatase) DUSP-10.
XX
XX Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;
KW neuronal degeneration syndrome; Alzheimer's disease; depression;
KW schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
KW osteoporosis; diabetes.
XX
XX Homo sapiens.
OS
XX MO200177340-A1.
PN
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-EP003966.
PF
XX 10-APR-2000; 2000EP-00107143.
PR
XX (MERE) MERCK PATENT GMBH.
PA
XX Duecker K;
PI
XX WPI; 2002-010917/01.
XX N-PSDB; AAS15768.
DR
XX Novel dual specificity phosphatase polypeptides useful for treating
PT cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
PT disease, depression, schizophrenia, asthma and immune disorders.
XX
XX Claim 2; Page 37-39; 43pp; English.
PS
XX The present invention relates to a new isolated dual specificity
XX phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
CC sequence that is fully defined in the specification. The invention also
CC provides a sequence encoded by a 3059 nucleotide sequence fully defined
CC in the specification, and a sequence having at least 95 % identity to the
CC polypeptide, or fragments or variants of DUSP-10. The invention is useful
CC for treating cancer e.g. leukemia, colon carcinoma, lung cancer,
CC prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
CC stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
CC schizophrenia, cardiac myopathies, asthma, immune disorders,
CC inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
CC osteoporosis, diabetes and diabetes associated diseases. The molecules of
CC the invention are also useful as vaccines for inducing immunological
CC response in a mammal, in disease diagnosis and in assays for screening
CC agonistic or antagonistic compounds. Other uses of the invention include
CC identifying membrane bound or soluble receptors, as a diagnostic reagent,
CC in chromosome localization studies, and as a valuable tool in tissue
CC expression studies. The present sequence represents the dual specificity
CC phosphatase, DUSP-10, protein of the invention
XX
XX Sequence 665 AA:
SQ
Query Match 99.2%; Score 1552; DB 5; Length 665;
Best Local Similarity 99.3%; Pred. No. 7,4e-163;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAHEMGTQVTRVALLESSTGTEKVLIDSRPFVEYNTSHILEAININSKMKRRLOQ 60
DB 1 MAHEMGTQVTRVALLESSTGTEKVLIDSRPFVEYNTSHILEAININSKMKRRLOQ 60
QY 61 DKYLITELIQSHAKHVIDSQKVVVYDQSSQDVASLSSDCLFTVLLGKLEKSPNSVHL 120
DB 61 DKYLITELIQSHAKHVIDSQKVVVYDQSSQDVASLSSDCLFTVLLGKLEKSPNSVHL 120
QY 121 LAGGPAFSPSCFPGCLCEGKSTLVPCTISOPCLPVPANIGPRLIPNLYLGQRPVNLKELI 180
DB 121 LAGGPAFSPSCFPGCLCEGKSTLVPCTISOPCLPVPANIGPRLIPNLYLGQRPVNLKELI 180
QY 181 QQNGIGVNLASYTCPEKDFIPESHFLRVVNDSPCEKILPWLDKSVDFIEKAKANGCV 240
DB 181 QQNGIGVNLASYTCPEKDFIPESHFLRVVNDSPCEKILPWLDKSVDFIEKAKANGCV 240

QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRPVKEKRPITISPNFNLGLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRPVKEKRPITISPNFNLGLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
RESULT 7
ID AAU75789 standard; protein; 665 AA.
XX AAU75789;
AC
XX
DT 08-MAY-2002 (first entry)
XX
XX Human protein phosphatase 7 (PP7) protein sequence.
DE
XX Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;
KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
KW melanoma; myeloma sarcoma.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Domain 11..131
FT 1/label= Rhodanese_like_domain
FT Region 15..170
FT 1/label= "Dual specificity protein phosphatase"
FT Region 85..298
FT 1/label= "Dual specificity protein phosphatase"
FT Domain 158..297
FT 1/label= Catalytic domain_DSPC
FT Region 220..281
FT 1/label= "Tyrosine specific protein phosphatases active
FT site tyr phosphatase.prif"
FT Region 242..254
FT 1/label= "Tyrosine phosphatase"
FT Region 242..252
FT 1/label= "Tyrosine specific protein phosphatases"
PN MO200210363-A2.
XX
XX 07-FEB-2002.
PD
XX
XX 26-JUL-2001; 2001WO-US023716.
PF
XX
XX 28-JUL-2000; 2000US-0221679P.
PR 03-AUG-2000; 2000US-0223372P.
PR 10-AUG-2000; 2000US-0224309P.
PR 18-AUG-2000; 2000US-0226728P.
PR 30-AUG-2000; 2000US-0229254P.
PR 08-SEP-2000; 2000US-0231366P.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;
PI Stewart EA, Gandhi AR, Patterson C, Lee EA, Hatalla AJA, Lu DAM;
PI Triboley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
PI Walla NK, Kearney L;
XX
XX WPI; 2002-188735/24.
DR N-PSDB; ABK14474.
XX
XX New protein phosphatases, useful for diagnosing, treating or preventing
PT immune system disorders (e.g. Crohn's disease), neurological disorders
PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.

PT cancers).

XX Claim 1; Page 107-108; 117pp; English.

CC The present invention relates to a new polypeptide, a naturally occurring
CC amino acid sequence at least 95 % identical to it, a biologically active
CC fragment of it or an immunogenic fragment of it. The polypeptides,
CC polynucleotides, agonists and antagonists are useful for diagnosing,
CC treating or preventing disorders associated with aberrant expression of
CC protein phosphatases (PP), particularly immune system disorders e.g.
CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
CC asthma or Crohn's disease, neurological disorders e.g. epilepsy,
CC Huntington's disease, dementia or Parkinson's disease, developmental
CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
CC or sarcoma. The present amino acid sequence represents human protein
CC phosphatase 7 (PP7) which is one of several human protein phosphatases
CC (AAU75783-AAU75792) of the invention

SQ Sequence 665 AA;

Query Match 99.2%; Score 1552; DB 5; Length 665;
Best Local Similarity 99.3%; Pred. No. 7,4e-163;

Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTEKRLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKIMKRLQ 60
DB 1 MAHEMIGTQVTEKRLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKIMKRLQ 60
QY 61 DKVLTITLIGHSAKHKVIDCSQKVVYDQSSQDVASLSDCFLTVLGLKLEKSFNSVHL 120
DB 61 DKVLTITLIGHSAKHKVIDCSQKVVYDQSSQDVASLSDCFLTVLGLKLEKSFNSVHL 120
QY 121 LAGFAFESRCFPGICBCKSTLPTCISQPCLPVANIPTRIIPNLVYLGCGQDVNLKELI 180
DB 121 LAGFAFESRCFPGICBCKSTLPTCISQPCLPVANIPTRIIPNLVYLGCGQDVNLKELI 180
QY 181 QONGIGVYLNASTCPKPDFIPESHFLRPVNDSCFKILPWLKSDVDFIEKAKASNGCV 240
DB 181 QONGIGVYLNASTCPKPDFIPESHFLRPVNDSCFKILPWLKSDVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIYIMKMDMSIDEAYRVVEKRPPTISPNFNLQQLDYEEKIKN 300
DB 241 LVHCLAGISRSATIAIYIMKMDMSIDEAYRVVEKRPPTISPNFNLQQLDYEEKIKN 300
QY 301 QT 302
DB 301 QT 302

RESULT 8

AB97946
ID ABB97946 standard; protein; 665 AA.

AC ABB97946;

DT 06-SEP-2002 (first entry)

DE Human protein sequence #13.

KW Human; brain; tonsil; hippocampus; foetal brain; diagnosis.

OS Homo sapiens.

PN WO200252005-A1.

PD 04-JUL-2002.

PF 20-DEC-2001; 2001WO-JP011217.

PR 22-DEC-2000; 2000JP-00389742.

PA (KAZU-) KAZUSA DNA RES INST FOUND.

PA (CELE-) CELESTAR LEXICO-SCI LTD.

PI Ohara O, Nagase T, Nakajima D;

DR WPI; 2002-500762/53.

DR N-PSDB; ABB93966.

PT Genes and their expression products cloned from human cDNA libraries for
treatment and diagnosis of diseases associated with their expression.

PS Claim 1(a); Page 112-116; 238pp; Japanese.

CC The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies
CC oligonucleotide probes and DNA chips for diagnosis and identification of
CC drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABB97934-ABB97964 represent
CC human proteins of the invention

SQ Sequence 665 AA;

Query Match 99.2%; Score 1552; DB 5; Length 665;
Best Local Similarity 99.3%; Pred. No. 7,4e-163;

Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTEKRLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKIMKRLQ 60
DB 1 MAHEMIGTQVTEKRLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKIMKRLQ 60
QY 61 DKVLTITLIGHSAKHKVIDCSQKVVYDQSSQDVASLSDCFLTVLGLKLEKSFNSVHL 120
DB 61 DKVLTITLIGHSAKHKVIDCSQKVVYDQSSQDVASLSDCFLTVLGLKLEKSFNSVHL 120
QY 121 LAGFAFESRCFPGICBCKSTLPTCISQPCLPVANIPTRIIPNLVYLGCGQDVNLKELI 180
DB 121 LAGFAFESRCFPGICBCKSTLPTCISQPCLPVANIPTRIIPNLVYLGCGQDVNLKELI 180
QY 181 QONGIGVYLNASTCPKPDFIPESHFLRPVNDSCFKILPWLKSDVDFIEKAKASNGCV 240
DB 181 QONGIGVYLNASTCPKPDFIPESHFLRPVNDSCFKILPWLKSDVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIYIMKMDMSIDEAYRVVEKRPPTISPNFNLQQLDYEEKIKN 300
DB 241 LVHCLAGISRSATIAIYIMKMDMSIDEAYRVVEKRPPTISPNFNLQQLDYEEKIKN 300
QY 301 QT 302
DB 301 QT 302

RESULT 9

AAU79929
ID AAU79929 standard; protein; 665 AA.

AC AAU79929;

DT 02-JUL-2002 (first entry)

DE Human dual specificity phosphatase 21117 protein.

OS Homo sapiens.

KW Human; dual specificity phosphatase 21117; erythroid-related disorder;
haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
erythrocytosis; liver-related disorder; cancer.

OS Homo sapiens.

Key Location/Qualifiers

FT Domain 11..131
/label= Rhodanese_like_domain

FT Region 21. .24
 FT /note= "Casein kinase II phosphorylation site"
 FT Region 91. .94
 FT /note= "Casein kinase II phosphorylation site"
 FT Domain 158. .297
 FT /note= "Dual specificity phosphatase catalytic domain"
 FT Region 214. .217
 FT /note= "Casein kinase II phosphorylation site"
 FT 242. .254
 FT /note= "Tyrosine specific protein phosphatase active site"
 FT 242. .254
 FT /note= "C-X-S-R motif"
 FT 266. .269
 FT /note= "Casein kinase II phosphorylation site"
 FT 369. .372
 FT /note= "Casein kinase II phosphorylation site"
 FT 421. .424
 FT /note= "Casein kinase II phosphorylation site"
 FT 434. .437
 FT /note= "Casein kinase II phosphorylation site"
 FT 458. .461
 FT /note= "Casein kinase II phosphorylation site"
 FT 508. .511
 FT /note= "Casein kinase II phosphorylation site"
 FT 589. .592
 FT /note= "Casein kinase II phosphorylation site"
 FT 612. .615
 FT /note= "Casein kinase II phosphorylation site"
 FT 617. .620
 FT /note= "Casein kinase II phosphorylation site"
 FT 642. .645
 FT /note= "Casein kinase II phosphorylation site"
 FT Region
 PN US002034807-A1.
 PD 21-MAR-2002.
 XX 23-MAR-2001, 2001US-00816494.
 XX 24-MAR-2000, 2000US-0191858P.
 PR
 XX (MEYE/) MEYERS R A.
 XX
 XX Meyers RA;
 PI
 XX WPI; 2002-351088/38.
 DR N-PSDB; ABK49402.
 XX
 XX New nucleic acids, designated 38692 and 21117, encoding dual specificity
 PT phosphatases for treating cell proliferation and differentiation
 PT disorders including hematopoietic and erythroid-related disorders and
 PT cancers.
 XX
 PS Claim 8; Fig 1; 76pp; English.
 CC The present invention relates to new nucleic acids designated 38692 and
 CC 21117 encoding dual specificity phosphatase family members. The nucleic
 CC acid, polypeptide encoded by it, and antibody specific for the
 CC polypeptide may be used to diagnose and treat haematopoietic-related
 CC disorders such as leukemias and autoimmune diseases, erythroid-related
 CC disorders such as anaemias and erythrocytosis, liver-related disorders,
 CC and cancers, particularly of the breast, colon, adipose, prostate and
 CC lung. The present amino acid sequence represents the human dual
 CC specificity phosphatase 21117 protein of the invention, as described
 CC above
 XX
 SO Sequence 665 AA:
 Query Match 99.2%; Score 1552; DB 5; Length 665;
 Best Local Similarity 99.3%; Pred. No. 7,4e-163;
 Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSGTEKYLIDSRPFVEYNTSHILEAININCKLMKRLQQ 60
 DB 1 MAHEMIGTQIVTERLVALLSGTEKYLIDSRPFVEYNTSHILEAININCKLMKRLQQ 60
 QY 61 DKVLITELIQSHAKKVIDCSQKVVVYDQSSQDVASLSDDCLTVLLGKLEKSFNSVHL 120
 DB 61 DKVLITELIQSHAKKVIDCSQKVVVYDQSSQDVASLSDDCLTVLLGKLEKSFNSVHL 120
 QY 121 LAGGFABFRCPCPGCLCEGKSTLVPCTISQPCFVANIQFRLIPNLYLGGQRLVNLKELI 180
 DB 121 LAGGFABFRCPCPGCLCEGKSTLVPCTISQPCFVANIQFRLIPNLYLGGQRLVNLKELI 180
 QY 181 QONGIGYVNASYTCRPPDIPESHFLRPVNDSPCEKILPMIDKSDVFTEKAKANGCV 240
 DB 181 QONGIGYVNASYTCRPPDIPESHFLRPVNDSPCEKILPMIDKSDVFTEKAKANGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFEKERTPTISNPNFLGQLDYEKKIKN 300
 DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFEKERTPTISNPNFLGQLDYEKKIKN 300
 QY 301 QT 302
 DB 301 QT 302
 DB 301 QT 302
 RESULT 10
 ABB97291
 ID ABB97291 standard; protein; 665 AA.
 XX
 XX ABB97291;
 AC
 XX
 DT 27-JUN-2002 (first entry)
 XX
 DE Novel human protein SEQ ID NO: 559.
 XX
 XX Human; antihaemic; vulnary; antiinflammatory; immunomodulator;
 KM antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KM neuroprotective; antiparkinsonian; protein therapy; EST;
 XX expressed sequence tag.
 XX
 OS Homo sapiens.
 OS
 XX
 PN MO200222660-A2.
 PD 21-MAR-2002.
 XX
 XX 10-SEP-2001; 2001WO-US026015.
 PF
 XX 11-SEP-2000; 2000US-00659671.
 PR
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QH, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX
 DR WPI; 2002-292408/33.
 DR N-PSDB; ABN32477.
 XX
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.
 XX
 PS Example 2; SEQ ID NO 559; 509pp; English.
 CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention

XX Sequence 665 AA;
SQ Query Match 99.2%; Score 1552; DB 5; Length 665;
Best Local Similarity 99.3%; Pred. No. 7,4e-163;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAHEMIGTQVTEVRLVALLSEGTETKVLIDSRPEVYENSHILEANINOSKIMKRLQO 60
DB 1 MAHEMIGTQVTEVRLVALLSEGTETKVLIDSRPEVYENSHILEANINOSKIMKRLQO 60
QY 61 DKVLTLLIQHSAGKVDIDCSQKVVVYDSSDVASLSDDCELVYLLGKLEKSPNSVHL 120
DB 61 DKVLTLLIQHSAGKVDIDCSQKVVVYDSSDVASLSDDCELVYLLGKLEKSPNSVHL 120
QY 121 LAGGFAFRRCPGLCEGKSTLVPTCISQPCLPVANIIGPRLILPNLYLGGQDVANKELI 180
DB 121 LAGGFAFRRCPGLCEGKSTLVPTCISQPCLPVANIIGPRLILPNLYLGGQDVANKELI 180
QY 181 QONGIGVYNASTCPEKDFIPESHFLRVVNDSPCEKILPMLDKSVDFIERAKASNGCV 240
DB 181 QONGIGVYNASTCPEKDFIPESHFLRVVNDSPCEKILPMLDKSVDFIERAKASNGCV 240
QY 241 LVHCLAGISRSATITAIYIMKRDMSLDRAVPRVKKRTTIPNRPVFLGQLIDYEKKIKN 300
DB 241 LVHCLAGISRSATITAIYIMKRDMSLDRAVPRVKKRTTIPNRPVFLGQLIDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
RESULT 11
AAB20325
ID AAB20325 standard; protein; 666 AA.
AC AAB20325;
XX 29-MAY-2001 (first entry)
DT 29-MAY-2001 (first entry)
XX
DE Human protein phosphatase and kinase protein-4.
XX
KW Protein phosphatase and kinase protein; PPHKP-4; human;
KW gastrointestinal disorder; immune system disorder; neurological disorder;
KW cell proliferative disorder; cancer; diagnosis; therapy.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 12 /note= "O-phosphorylated"
FT Modified-site 21 /note= "O-phosphorylated"
FT Modified-site 23 /note= "O-phosphorylated"
FT Modified-site 38 /note= "O-phosphorylated"
FT Modified-site 38 /note= "N-glycosylated"
FT Modified-site 38 /note= "N-glycosylated"
FT Modified-site 38 /note= "N-glycosylated"
FT Modified-site 49 /note= "N-glycosylated"
FT Modified-site 72 /note= "N-glycosylated"
FT Modified-site 82 /note= "O-phosphorylated"
FT Modified-site 82 /note= "O-phosphorylated"
FT Region 85..298 /note= "VH1-type dual specificity phosphatase signature"
FT Modified-site 91 /note= "O-phosphorylated"
FT Modified-site 190 /note= "O-phosphorylated"
FT Modified-site 212 /note= "N-glycosylated"
FT Modified-site 212 /note= "N-glycosylated"

FT Modified-site 214 /note= "O-phosphorylated"
FT Active-site 220..280 /note= "tyrosine specific protein phosphatase"
FT Region 237..278 /note= "Y phosphatase signature"
FT Modified-site 266 /note= "O-phosphorylated"
FT Modified-site 280 /note= "O-phosphorylated"
FT Modified-site 300 /note= "O-phosphorylated"
FT Modified-site 369 /note= "N-glycosylated"
FT Modified-site 393 /note= "O-phosphorylated"
FT Modified-site 421 /note= "O-phosphorylated"
FT Modified-site 422 /note= "O-phosphorylated"
FT Modified-site 434 /note= "O-phosphorylated"
FT Modified-site 439 /note= "O-phosphorylated"
FT Modified-site 468 /note= "O-phosphorylated"
FT Modified-site 471 /note= "O-phosphorylated"
FT Modified-site 479 /note= "O-phosphorylated"
FT Modified-site 528 /note= "O-phosphorylated"
FT Modified-site 590 /note= "O-phosphorylated"
FT Modified-site 597 /note= "O-phosphorylated"
FT Modified-site 605 /note= "O-phosphorylated"
FT Modified-site 610 /note= "O-phosphorylated"
FT Modified-site 613 /note= "O-phosphorylated"
FT Modified-site 618 /note= "O-phosphorylated"
FT Modified-site 628 /note= "O-phosphorylated"
FT Modified-site 641 /note= "O-phosphorylated"
FT Modified-site 643 /note= "N-glycosylated"
FT Modified-site 643 /note= "O-phosphorylated"
XX
XX WO200120004-A2.
XX
XX 22-MAR-2001.
XX
XX 14-SEP-2000; 2000WO-US025515.
XX
XX 15-SEP-1999; 99US-015414P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y;
XX Lu DM;
XX WPI; 2001-244811/25.
XX
XX N-PSDB; AAF30479.
XX
XX Novel human protein phosphatase and kinase proteins for diagnosis,
XX treatment and prevention of gastrointestinal, immune system, neurological
XX and cell proliferative disorders.
XX
XX Claim 1; Page 87-88; 103pp; English.
XX

CC The present sequence is that of novel human protein phosphatase and
CC kinase protein PPHK-4, as predicted from Incyte Clone ID No. 1234795CB1
CC (see Aa30479). Tissues that express PPHK-4 (as a fraction of total
CC tissues expressing PPHK-4) include gastrointestinal (0.365),
CC cardiovascular (0.154), nervous (0.154) and reproductive (0.154).
CC Diseases or conditions associated with tissues expressing PPHK-4 (as a
CC fraction of total tissues expressing PPHK-4) include cancer (0.692),
CC inflammation or trauma (0.308) and cell proliferation (0.231). The
CC encoded protein shows homology to mouse neuronal cytosine threonine
CC phosphatase 1. The invention provides human PPHK-1 to -11 polypeptides
CC (see Aa20322-32) and polynucleotides (see AaF30476-86). It also provides
CC expression vectors, host cells, antibodies, agonists and antagonists, as
CC well as methods for diagnosing, treating or preventing disorders
CC associated with expression of PPHK, including gastrointestinal
CC disorders, immune system disorders, neurological disorders and cell
CC proliferative disorders, including cancer

SQ Sequence 666 AA;

Query Match 99.2%; Score 1552; DB 4; Length 666;
Best Local Similarity 99.3%; Pred. No. 7.4e-163; Indels 0; Gaps 0;
Matches 300; Conservative 1; Mismatches 1;

QY 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
QY 61 DKVLITELIQHSARKKVDIDCSQKVVYVYDSSQDVASLSDDCFLYVLGLKLEKSFNSVHL 120
DB 61 DKVLITELIQHSARKKVDIDCSQKVVYVYDSSQDVASLSDDCFLYVLGLKLEKSFNSVHL 120
QY 121 LAGGFABFRCFPGCLCEGKSTLVPTCISOPCLPVANIGPTRLIPNLVYLGQGRVNLKELI 180
DB 121 LAGGFABFRCFPGCLCEGKSTLVPTCISOPCLPVANIGPTRLIPNLVYLGQGRVNLKELI 180
QY 181 QONGIGYVNAASYTCRPPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
DB 181 QONGIGYVNAASYTCRPPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

RESULT 12

ADA54744

ID ADA54744 standard; protein; 665 AA.

AC ADA54744;

DT 20-NOV-2003 (first entry)

DE Human protein, SEQ ID 2312.

KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nocrotropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.

OS Homo sapiens.

PN EP1293569-A2.

PD 19-MAR-2003.

PF 21-MAR-2002; 2002EP-0006586.

PR 14-SEP-2001; 2001JP-0032831.
PR 24-JAN-2002; 2002US-0350435P.

XX

PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto U, Isono Y, Hio Y, Otsuka K, Nagai K, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;
XX WPI: 2003-395539/38.
DR N-PSDB; ADA53105.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.

PS Claim 14; SEQ ID NO 2312; 205pp; English.

CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA5433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.

SQ Sequence 665 AA;

Query Match 98.8%; Score 1545; DB 6; Length 665;
Best Local Similarity 99.0%; Pred. No. 4.4e-162; Indels 0; Gaps 0;
Matches 299; Conservative 1; Mismatches 2;

QY 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
QY 61 DKVLITELIQHSARKKVDIDCSQKVVYVYDSSQDVASLSDDCFLYVLGLKLEKSFNSVHL 120
DB 61 DKVLITELIQHSARKKVDIDCSQKVVYVYDSSQDVASLSDDCFLYVLGLKLEKSFNSVHL 120
QY 121 LAGGFABFRCFPGCLCEGKSTLVPTCISOPCLPVANIGPTRLIPNLVYLGQGRVNLKELI 180
DB 121 LAGGFABFRCFPGCLCEGKSTLVPTCISOPCLPVANIGPTRLIPNLVYLGQGRVNLKELI 180
QY 181 QONGIGYVNAASYTCRPPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
DB 181 QONGIGYVNAASYTCRPPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

RESULT 13

AAU79161

ID AAU79161 standard; protein; 665 AA.

AC AAU79161;

DT 02-JUL-2002 (first entry)

DE Human dual-specificity phosphatase-3 (DSP-16) mutant protein #1.

KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW cell cycle; abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; mutant; mutein.

OS Homo sapiens.

OS Synthetic.

XX

RH	Key	Location/Qualifiers
FT	Misc-difference	213 /note= "Wild-type Asp substituted by Ala"
XX		
XX		
PV	MO200226997-A2.	
XX		
PD	04-APR-2002.	
XX		
PF	25-SEP-2001; 2001WO-US030124.	
XX		
PR	26-SEP-2000; 2000US-0235487P.	
XX		
PA	(CEPT-) CEPTYR INC.	
PI	Luche RM, Wei B;	
DR	WPI; 2002-315802/35.	
XX		
PT	New DSP-16 polypeptide, useful for identifying modulators of its activity, which can be used in the treatment of disorders such as Duchenne muscular dystrophy, or cancer.	
XX		
PS	Claim 46; Page; 87pp; English.	
CC	The present invention relates to a new polypeptide, DSP-16, having a 665 amino acid sequence, given in the specification, or a variant having at least 50 % identical residues, which retains the ability to dephosphorylate an activated mitogen-activated protein (MAP) kinase. The CC	
CC	activity, for modulation of a proliferative response in a cell, survival of a cell, or differentiation of a cell. The cell displays contact inhibition of cell growth or anchorage independent growth and may display altered intercellular adhesion. The agent may modulate apoptosis, or the cell cycle. The identified modulators can be used to treat Duchenne CC	
CC	muscular dystrophy, cancer, graft-versus-host disease, autoimmune diseases, allergies, metabolic diseases, abnormal cell growth, abnormal CC	
CC	cell proliferation, and cell cycle abnormalities. The present amino acid sequence represents the human dual-specificity phosphatase-3 (DSP-16) mutant protein #1. Note: This sequence is not shown in the specification CC	
CC	but is derived from the wild-type human DSP-16 (AAU79156) protein given in figure 2 of the specification	
CC		
XX		
SQ	Sequence 665 AA;	
	Query Match 98.7%; Score 1544; DB 5; Length 665;	
	Best Local Similarity 99.0%; Pred. No. 5,7e-162;	
	Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0	
QY	1 MAHEMIGTIVTERLVALLSEGTCKLLIDSRPFVEYNNSHLEAININCSKLMRRLOO 60	
DB	1 MAHEMIGTIVTERLVALLSEGTCKLLIDSRPFVEYNNSHLEAININCSKLMRRLOO 60	
QY	61 DKVLITTELIQSHAKKHVDIDCSQKVVVYQSSODVASLSDFCLFYLIAKLEKSFNVSVL 120	
DB	61 DKVLITTELIQSHAKKHVDIDCSQKVVVYQSSODVASLSDFCLFYLIAKLEKSFNVSVL 120	
QY	121 IAGGFAEBSRCPEGLCEGKSTLVPTCISQCLPVANIGETRIIPNLVILGCQGDVINKEII 180	
DB	121 IAGGFAEBSRCPEGLCEGKSTLVPTCISQCLPVANIGETRIIPNLVILGCQGDVINKEII 180	
QY	181 QONGIGVYLNASVTCPKDPFIESHRLRPVNDSPCEKLLPMIDKSVDTIEKAKSNGCV 240	
DB	181 QONGIGVYLNASVTCPKDPFIESHRLRPVNDSPCEKLLPMIDKSVDTIEKAKSNGCV 240	
QY	241 LVHCAGISRSAATIAIAYIMKMDMSLDAAYRFVKEKERPTISNFNPLSQLLDYEEKIKN 300	
DB	241 LVHCAGISRSAATIAIAYIMKMDMSLDAAYRFVKEKERPTISNFNPLSQLLDYEEKIKN 300	
QY	301 QT 302	
DB	301 QT 302	

RESULT 14

ABRS2352
ID ABR52352 standard; protein, 665 AA.
XX
XX ABR52352;
XX
XX ABR52352;
DT 19-JUN-2003 (first entry)
XX
DE Protein relating to the invention SEQ ID NO: 42.
XX
XX antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antiproliferative; cardiac; cytostatic; gene therapy; liver disease;
XX antiproliferative disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
XX Homo sapiens.
OS
XX WO200257460-A2.
PN
XX 25-JUL-2002.
PD
XX
XX 20-DEC-2001; 2001WO-US050459.
PF
XX
XX 20-DEC-2000; 2000US-0256868P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
PA
XX Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Baasolino D;
PI Krystek S, Mcatee P, Suchard S, Banas D;
XX
XX WPI; 2002-599721/64.
DR N-PSDB; ACC60521.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
XX Claim 5; Fig 5; 801PP; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antiarthritic, antipsoriatic, cardiac, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX
XX
SQ Sequence 665 AA;

Query Match 98.7%; Score 1544; DB 5; Length 665;
Best Local Similarity 99.0%; Pred. No. 5,7e-162;
Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0

1 MAHEMIGQIWTERRVALLSEGTETKLLIDSRPFVYNTSHILEAININCSKLMRRRLQQ 60
1 MAHEMIGQIWTERRVALLSEGTETKLLIDSRPFVYNTSHILEAININCSKLMRRRLQQ 60

61 DKVLTLEIIGHSAKKRVIDDSQKRVYVDQSSQVYASLSSDCEFLVVLGKLEKSPSYHL 120
61 DKVLTLEIIGHSAKKRVIDDSQKRVYVDQSSQVYASLSSDCEFLVVLGKLEKSPSYHL 120

QY 121 LAGGFAEFRCFPGLCGKSTLVPCTISQPCLPVANIGPRLIPNLVYLGQRVNLKEL 180
 DB 121 LAGGFAEFRCFPGLCGKSTLVPCTISQPCLPVANIGPRLIPNLVYLGQRVNLKEL 180
 QY 181 QONGIGVYNASTYCPKDPFIPESHFLRVVNDSPCEKILPMDKSVDFIEKKAANGCV 240
 DB 181 QONGIGVYNASTYCPKDPFIPESHFLRVVNDSPCEKILPMDKSVDFIEKKAANGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPITSPNFVLGQLDYEKKIN 300
 DB 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPITSPNFVLGQLDYEKKIN 300
 QY 301 QT 302
 DB 301 QT 302

RESULT 15
 AAM25744
 ID AAM25744 standard; protein; 672 AA.
 XX
 AC AAM25744;
 DT 16-OCT-2001 (first entry)
 DE Human protein sequence SEQ ID NO:1259.
 XX
 KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiant; central nervous system; virucide;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KW antileggregant; haemostatic; vulnery; antilicer; osteopathic; eczema;
 KW dermatological; antiallergic; antidiabetic; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW chromocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder.
 XX
 OS Homo sapiens.
 XX
 PN W0200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 22-DEC-2000; 2000MO-US035017.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-457603/49.
 DR N-PSDB; AAM99685.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
 XX
 PS Claim 20; Page 260; 1217pp; English.
 XX
 CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and cells
 CC they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;

CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
 CC antilicer; osteopathic; dermatological; antiarthritic; antidiabetic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic disease, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders
 XX
 SQ Sequence 672 AA;

Query Match 98.7%; Score 1544; DB 4; Length 672;
 Best Local Similarity 99.0%; Fred. No. 5.8e-162;
 Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY	1	MAHEMIGTOIVTERLVALLESGETEKLIDSRPFVEYNTSHLEAININCSKIMKRLQ	60
DB	8	MAHEMIGTOIVTERGVALLESGETEKLIDSRPFVEYNTSHLEAININCSKIMKRLQ	67
QY	61	DKVLITELIQHSKHKVDIDSGKVYVYDQSSQDVASLSSDCELVYLGKLEKSFNSVHL	120
DB	68	DKVLITELIQHSKHKVDIDSGKVYVYDQSSQDVASLSSDCELVYLGKLEKSFNSVHL	127
QY	121	LAGGFAEFRCFPGLCGKSTLVPCTISQPCLPVANIGPRLIPNLVYLGQRVNLKEL	180
DB	128	LAGGFAEFRCFPGLCGKSTLVPCTISQPCLPVANIGPRLIPNLVYLGQRVNLKEL	187
QY	181	QONGIGVYNASTYCPKDPFIPESHFLRVVNDSPCEKILPMDKSVDFIEKKAANGCV	240
DB	188	QONGIGVYNASTYCPKDPFIPESHFLRVVNDSPCEKILPMDKSVDFIEKKAANGCV	247
QY	241	LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPITSPNFVLGQLDYEKKIN	300
DB	248	LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPITSPNFVLGQLDYEKKIN	307
QY	301	QT 302	
DB	308	QT 309	

Search completed: June 21, 2004, 13:20:33
 Job time : 33.7291 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:19:48 ; Search time 25.9214 Seconds

(Without alignment)
3289.124 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302

Perfect score: 1564
Sequence: 1 MAHEMGTQIVTERLVALLE.....PNNFPGQLDYEKIKNT 302Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 28213646 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubppa/US10A_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US10B_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1552	99.2	662 12 US-10-072-012-258	Sequence 258, App
2	1552	99.2	665 9 US-09-816-494-2	Sequence 2, Appl
3	1552	99.2	665 9 US-09-964-277-2	Sequence 2, Appl
4	1552	99.2	665 12 US-10-072-012-680	Sequence 680, App
5	1552	99.2	665 12 US-10-168-506-14	Sequence 14, Appl
6	1552	99.2	665 12 US-10-343-357-7	Sequence 7, Appl
7	1552	99.2	665 15 US-10-377-072-26	Sequence 26, Appl
8	1552	99.2	665 15 US-10-257-026-2	Sequence 2, Appl
9	1552	99.2	665 16 US-10-648-593-240	Sequence 240, App
10	1552	99.2	665 16 US-10-648-593-247	Sequence 247, App
11	1552	99.2	690 12 US-10-072-012-679	Sequence 679, App
12	1552	99.2	690 12 US-10-072-012-703	Sequence 703, App
13	1552	99.2	690 12 US-10-825-114-54204	Sequence 54204, A
14	1545	98.8	665 12 US-10-072-012-681	Sequence 681, App
15	1545	98.8	665 15 US-10-094-749-2312	Sequence 2312, App

16	1544	98.7	672 12 US-10-296-115-1259	Sequence 1259, App
17	1517	97.0	660 12 US-10-072-012-682	Sequence 682, App
18	1514	96.8	677 12 US-10-072-012-683	Sequence 683, App
19	1504.5	96.2	680 12 US-10-072-012-256	Sequence 256, App
20	988	63.2	625 12 US-10-072-012-699	Sequence 699, App
21	985	63.0	663 12 US-10-072-012-700	Sequence 700, App
22	837	53.5	616 12 US-10-072-012-266	Sequence 266, App
23	807	51.6	155 9 US-09-964-277-7	Sequence 7, Appl
24	654.5	41.8	169 14 US-10-346-356-15	Sequence 15, Appl
25	654.5	41.8	170 9 US-09-775-925-26	Sequence 26, Appl
26	654.5	41.8	170 9 US-09-847-519A-11	Sequence 11, Appl
27	654.5	41.8	170 12 US-10-655-073-16	Sequence 16, Appl
28	654.5	41.8	170 14 US-10-314-058-14	Sequence 14, Appl
29	654.5	41.8	170 14 US-10-405-808-16	Sequence 16, Appl
30	646	41.3	517 9 US-09-964-277-21	Sequence 21, Appl
31	631.5	40.4	461 12 US-10-072-012-701	Sequence 701, App
32	620	39.6	501 12 US-10-072-012-702	Sequence 702, App
33	619	39.6	155 9 US-09-964-277-6	Sequence 6, Appl
34	619	39.6	155 9 US-09-955-732-6	Sequence 6, Appl
35	469	30.0	444 9 US-09-964-899-47	Sequence 47, Appl
36	469	30.0	482 12 US-10-058-270A-130	Sequence 130, App
37	469	30.0	482 14 US-10-346-356-2	Sequence 2, Appl
38	469	30.0	482 16 US-10-648-593-164	Sequence 164, App
39	462.5	29.6	381 14 US-10-184-832-2	Sequence 2, Appl
40	426	27.2	394 12 US-10-060-065-25	Sequence 25, Appl
41	426	27.2	394 14 US-10-059-585-46	Sequence 46, Appl
42	425	27.2	394 9 US-09-736-457-805	Sequence 805, App
43	425	27.2	394 9 US-09-736-457-827	Sequence 827, App
44	425	27.2	394 9 US-09-902-941-805	Sequence 805, App
45	425	27.2	394 9 US-09-902-941-827	Sequence 827, App

ALIGNMENTS

RESULT 1
US-10-072-012-258
; Sequence 258, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchenev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Smitkeet, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Futrak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072, 012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265, 102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265, 514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265, 517
; PRIOR FILING DATE: 2001-01-31

```

; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 258
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-258
```

```

Query Match          99.2%; Score 1552; DB 12; Length 662;
Best Local Similarity 99.3%; Pred. No. 3.4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIHSAAKHKVDIDCSQKVVVYDSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIHSAAKHKVDIDCSQKVVVYDSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIGPTTILPNTLYLGCCRDVNLKELI 180
DB 121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIGPTTILPNTLYLGCCRDVNLKELI 180
QY 181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPWLKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPWLKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPITISPENFLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPITISPENFLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
```

```

RESULT 2
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-494-2
```

```

Query Match          99.2%; Score 1552; DB 9; Length 665;
Best Local Similarity 99.3%; Pred. No. 3.4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIHSAAKHKVDIDCSQKVVVYDSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIHSAAKHKVDIDCSQKVVVYDSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIGPTTILPNTLYLGCCRDVNLKELI 180
DB 121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIGPTTILPNTLYLGCCRDVNLKELI 180
QY 181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPWLKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPWLKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPITISPENFLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPITISPENFLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
```

```

RESULT 3
```

```

US-09-964-277-2
```

```

; Sequence 2, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125, 434
; CURRENT APPLICATION NUMBER: US/09/964,277
; PRIOR FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-964-277-2
```

```

Query Match          99.2%; Score 1552; DB 9; Length 665;
Best Local Similarity 99.3%; Pred. No. 3.4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIHSAAKHKVDIDCSQKVVVYDSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIHSAAKHKVDIDCSQKVVVYDSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIGPTTILPNTLYLGCCRDVNLKELI 180
DB 121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIGPTTILPNTLYLGCCRDVNLKELI 180
QY 181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPWLKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPWLKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPITISPENFLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPITISPENFLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
```

Db 301 QT 302

RESULT 4

US-10-072-012-680
; Sequence 680, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patrujan, Meera
; APPLICANT: Shinkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Baha
; APPLICANT: Padigaru, Muraidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-680

Query Match 99.2%; Score 1552; DB 12; Length 665;
Best Local Similarity 99.3%; Pred. No 3, 4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
DB 1 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60

QY 61 DKVLITELIQHSAKHKVDIDCSQKVVYDQSSQDVASLSDDCFLLVLLGLKLEKSFNSVHL 120
DB 61 DKVLITELIQHSAKHKVDIDCSQKVVYDQSSQDVASLSDDCFLLVLLGLKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIGPTRILPNLYLGGQDVLNKELI 180
DB 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIGPTRILPNLYLGGQDVLNKELI 180
QY 181 QONGIGYVLANASTYCKPDPFIPESHFLRVVNDSPCEKILPMWLDKSVDFLEKAKANGCV 240
DB 181 QONGIGYVLANASTYCKPDPFIPESHFLRVVNDSPCEKILPMWLDKSVDFLEKAKANGCV 240
QY 241 LVHCLAGISRSATIALAYIMKQWMSLDEAYRFVKEKRPDISPBNFLGQLDYEKKIKX 300
DB 241 LVHCLAGISRSATIALAYIMKQWMSLDEAYRFVKEKRPDISPBNFLGQLDYEKKIKX 300
QY 301 QT 302
DB 301 QT 302

RESULT 5

US-10-168-506-14
; Sequence 14, Application US/10168506
; Publication No. US20040053229A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERRARD
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: HILL, RON
; APPLICANT: PLANAGAN, BETTER
; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
; FILE REFERENCE: 038602/1351
; CURRENT APPLICATION NUMBER: US/10/168,506
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: PCT/US00/34736
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-168-506-14

Query Match 99.2%; Score 1552; DB 12; Length 665;
Best Local Similarity 99.3%; Pred. No 3, 4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
DB 1 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
QY 61 DKVLITELIQHSAKHKVDIDCSQKVVYDQSSQDVASLSDDCFLLVLLGLKLEKSFNSVHL 120
DB 61 DKVLITELIQHSAKHKVDIDCSQKVVYDQSSQDVASLSDDCFLLVLLGLKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIGPTRILPNLYLGGQDVLNKELI 180
DB 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIGPTRILPNLYLGGQDVLNKELI 180
QY 181 QONGIGYVLANASTYCKPDPFIPESHFLRVVNDSPCEKILPMWLDKSVDFLEKAKANGCV 240
DB 181 QONGIGYVLANASTYCKPDPFIPESHFLRVVNDSPCEKILPMWLDKSVDFLEKAKANGCV 240
QY 241 LVHCLAGISRSATIALAYIMKQWMSLDEAYRFVKEKRPDISPBNFLGQLDYEKKIKX 300
DB 241 LVHCLAGISRSATIALAYIMKQWMSLDEAYRFVKEKRPDISPBNFLGQLDYEKKIKX 300
QY 301 QT 302
DB 301 QT 302

Db 301 QT 302

RESULT 6

US-10-343-357-7
; Sequence 7, Application US/10343357
; Publication No. US20040058341A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
; APPLICANT: ELIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
; APPLICANT: LEE, Ernestine A.; HAFALIA, April J. A.
; APPLICANT: LU, Dying Aina M.; TRIBOULEY, Catherine M.
; APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
; APPLICANT: YUE, Henry; WARREN, Bridget A.
; APPLICANT: NGUYEN, Daniel B.; CHAMLA, Nandinder K.
; APPLICANT: KEARNEY, Liam
; TITLE OF INVENTION: PROTEIN PHOSPHATASES
; FILE REFERENCE: PI-0173 PCT
; CURRENT APPLICATION NUMBER: US/10/343,357
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: PCT/US01/23716
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,679
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/223,272
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/224,309
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 60/226,728
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/229,254
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: US 60/231,366
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CD1
US-10-343-357-7

Query Match 99.2%; Score 1552; DB 12; Length 665;
Best Local Similarity 99.3%; Pred. No. 3.4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTOYTERVALLESGETEKVLLIDSRPFVYNTSHILEAININCSKIMRRLOQ 60
DB 1 MAHEMIGTOYTERVALLESGETEKVLLIDSRPFVYNTSHILEAININCSKIMRRLOQ 60
QY 61 DKVLTIELIHSASHAKHVIDIDCSQKVVVYDQSSQVVASLSSDCFLVYLLGKLEKSFNSVHL 120
DB 61 DKVLTIELIHSASHAKHVIDIDCSQKVVVYDQSSQVVASLSSDCFLVYLLGKLEKSFNSVHL 120
QY 121 LAGFAEFSCFCFGLCEGSKSTLVPTCTISQCLPVANIGPTRILPMLYLGCGQRDVLANKEIM 180
DB 121 LAGFAEFSCFCFGLCEGSKSTLVPTCTISQCLPVANIGPTRILPMLYLGCGQRDVLANKEIM 180
QY 181 QONGIGYVLNASYTCPPDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIERAKASNGCV 240
DB 181 QONGIGYVLNASYTCPPDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIERAKASNGCV 240
QY 241 LVHCLAGISSATTAIAYIMKRMDSIDEAYRFVKEKRPITISPNFNLGQLLDYEEKIKN 300
DB 241 LVHCLAGISSATTAIAYIMKRMDSIDEAYRFVKEKRPITISPNFNLGQLLDYEEKIKN 300
QY 301 QT 302

Db 301 QT 302

RESULT 7

US-10-377-072-26
; Sequence 26, Application US/10377072
; Publication No. US2004009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Miyoung
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: MPI03-0180NMIM
; CURRENT APPLICATION NUMBER: US/10/377,072
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-377-072-26

Query Match 99.2%; Score 1552; DB 15; Length 665;
Best Local Similarity 99.3%; Pred. No. 3.4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTOYTERVALLESGETEKVLLIDSRPFVYNTSHILEAININCSKIMRRLOQ 60
DB 1 MAHEMIGTOYTERVALLESGETEKVLLIDSRPFVYNTSHILEAININCSKIMRRLOQ 60
QY 61 DKVLTIELIHSASHAKHVIDIDCSQKVVVYDQSSQVVASLSSDCFLVYLLGKLEKSFNSVHL 120
DB 61 DKVLTIELIHSASHAKHVIDIDCSQKVVVYDQSSQVVASLSSDCFLVYLLGKLEKSFNSVHL 120
QY 121 LAGFAEFSCFCFGLCEGSKSTLVPTCTISQCLPVANIGPTRILPMLYLGCGQRDVLANKEIM 180
DB 121 LAGFAEFSCFCFGLCEGSKSTLVPTCTISQCLPVANIGPTRILPMLYLGCGQRDVLANKEIM 180
QY 181 QONGIGYVLNASYTCPPDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIERAKASNGCV 240
DB 181 QONGIGYVLNASYTCPPDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIERAKASNGCV 240

QY	241	IYHCLAGISRSATIAIAYIMKMDMSLDEAYR	FVKEKPTISPNPNF	IGQLDYEKKIKN	300
Db	241	IYHCLAGISRSATIAIAYIMKMDMSLDEAYR	FVKEKPTISPNPNF	IGQLDYEKKIKN	300
QY	301	QT	302		
Db	301	QT	302		

```

RESULT 8
US-10-257-026-2
; Sequence 2, Application US/10257026
; Publication NO. US20040086659A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10XOMS
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-257-026-2

```

Query Match	99.2%	Score 1552;	DB 16;	Length 665;
Best Local Similarity	99.3%	Pred. No. 3.4e-147;		
Matches 300; Conservative	1;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	1	MAHEMGTQIVTERVALLAESGTEKVLIDSNPEVNTSHLEAININSCIKMKRRLQ	60
Db	1	MAHEMGTQIVTERVALLAESGTEKVLIDSNPEVNTSHLEAININSCIKMKRRLQ	60
Qy	1	MAHEMGTQIVTERVALLAESGTEKVLIDSNPEVNTSHLEAININSCIKMKRRLQ	60
Db	1	MAHEMGTQIVTERVALLAESGTEKVLIDSNPEVNTSHLEAININSCIKMKRRLQ	60
Qy	61	DKLTITELI,OHSAKHVYDIDGOKRVVYDQSSODVASISDPCFLTITLLGKLEKSPNSVHL	120
Db	61	DKLTITELI,OHSAKHVYDIDGOKRVVYDQSSODVASISDPCFLTITLLGKLEKSPNSVHL	120
Qy	121	LAGGFAEFRCFPGICEGKSTLVPTCISQPCLPVANIPTRIIPNLVYGCQRDVYLNKELI	180
Db	121	LAGGFAEFRCFPGICEGKSTLVPTCISQPCLPVANIPTRIIPNLVYGCQRDVYLNKELI	180
Qy	121	LAGGFAEFRCFPGICEGKSTLVPTCISQPCLPVANIPTRIIPNLVYGCQRDVYLNKELI	180
Db	121	LAGGFAEFRCFPGICEGKSTLVPTCISQPCLPVANIPTRIIPNLVYGCQRDVYLNKELI	180
Qy	181	QONGIGVYLNASTYCKPDPIPESHFLRPVNDSCFKILPMLKSDVTEKAAASNGCV	240
Db	181	QONGIGVYLNASTYCKPDPIPESHFLRPVNDSCFKILPMLKSDVTEKAAASNGCV	240
Qy	241	LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVVEKRPPTISPNPNFLGQLLDYEKKIKN	300
Db	241	LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVVEKRPPTISPNPNFLGQLLDYEKKIKN	300
Qy	301	QT 302	
Db	301	QT 302	

```

RESULT 9
US-10-648-593-240
? Sequence 240, Application US/10648593
? Publication No. US20040106132a1
? GENERAL INFORMATION:
? APPLICANT: Bristol-Myers Squibb Company
? TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
? TITLE OR INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
? TITLE OR INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
? FILE REFERENCE: D0273 NP
? CURRENT APPLICATION NUMBER: US/10/648,593
? CURRENT FILING DATE: 2003-08-26
? PRIOR APPLICATION NUMBER: 60/406,385
? PRIOR FILING DATE: 2002-08-27
? NUMBER OF SEQ ID NOS: 557
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 240

```

```

; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-240

```

Query Match	99.2%	Score 1552;	DB 16;	Length 665;
Best Local Similarity	99.3%	Pred. No. 3,4e-147;		
Matches 300; Conservative	1;	Mismatches	1;	Indels 0; Gaps 0;

Qy	1	MAHEMGTQIVTERVVALLESCTEKVLLIDSRPVEVYNTSHILEAININCSKLMKRLQ	60
Qy	1	MAHEMGTQIVTERVVALLESCTEKVLLIDSRPVEVYNTSHILEAININCSKLMKRLQ	60
Db	1	MAHEMGTQIVTERVVALLESCTEKVLLIDSRPVEVYNTSHILEAININCSKLMKRLQ	60
Qy	61	DKLITLLELIQHSAKKKVIDCSQKRVVYDSSQVVASLSSPCFLTULLGLTLESFNSVHL	120
Db	61	DKLITLLELIQHSAKKKVIDCSQKRVVYDSSQVVASLSSPCFLTULLGLTLESFNSVHL	120
Qy	121	LAGFAFESRCPGLCEGKSTLVPTCSQPLPAANIGPTRILPNLYLGCQRVLNKLKI	180
Db	121	LAGFAFESRCPGLCEGKSTLVPTCSQPLPAANIGPTRILPNLYLGCQRVLNKLKI	180
Qy	181	QONGIGYVLNASYTCPKPDFIPESHFLFVPVNDSPCEKILPMLDKSYDFIEKAKASNGCV	240
Db	181	QONGIGYVLNASYTCPKPDFIPESHFLFVPVNDSPCEKILPMLDKSYDFIEKAKASNGCV	240
Qy	241	LVHCLAGISRSATTAIAIYIMKRMDSLDEAVRFYKXRPRTISPNFNLGQLLYEKKIKN	300
Db	241	LVHCLAGISRSATTAIAIYIMKRMDSLDEAVRFYKXRPRTISPNFNLGQLLYEKKIKN	300

```

RESULT 10
US-10-648-593-247
: Sequence 247, Application US/10648593
: Publication No. US20040106132A1
: GENERAL INFORMATION:
: APPLICANT: Bristol-Myers Squibb Company
: TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
: TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
: TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
: FILE REFERENCE: D0273 NP
: CURRENT APPLICATION NUMBER: US/10/648,593
: CURRENT FILING DATE: 2003-08-26
: PRIOR APPLICATION NUMBER: 60/406,385
: PRIOR FILING DATE: 2002-08-27
: NUMBER OF SEQ ID NOS: 557
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 247
: LENGTH: 665
: TYPE: prt
: ORGANISM: Homo sapiens
: US-10-648-593-247

```

	Query Match	Similarity	Score	DB	Length
Best Local	99.24	99.34	1552	16	665
Matches	300	Conservative	1	Mismatches	1
				Indels	0
				Gaps	0

Db 161 QONGIGYVLMASNTCPKDPFIPESHFLRVPVNDSPCEKILPWLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATITAIYIMKRMDSLDEAYRFVKEKPTTISPNFNLGQLDYEKKIKN 300
Db 241 LVHCLAGISRSATITAIYIMKRMDSLDEAYRFVKEKPTTISPNFNLGQLDYEKKIKN 300
QY 301 QT 302
Db 301 QT 302

RESULT 11
US-10-072-012-679
Sequence 679, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tcherenev, Vellizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernhusen, Bryan
APPLICANT: Patuturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Beha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072.012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 679
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-679

Query Match 99.2%; Score 1552; DB 12; Length 690;
Best Local Similarity 99.3%; Pred. No. 3.6e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAHMTGTQVTERLVALLSSGTBKVLLIDSRPVEYNTSHIIEAININISKIMKRLQ 60
Db 26 MAHMTGTQVTERLVALLSSGTBKVLLIDSRPVEYNTSHIIEAININISKIMKRLQ 85
QY 61 DKVLTTELIOHSAKHVVDICSOQVWVYDOSQDVASLSSDCFLTYLLGLKLESPNSVHL 120
Db 86 DKVLTTELIOHSAKHVVDICSOQVWVYDOSQDVASLSSDCFLTYLLGLKLESPNSVHL 145
QY 121 LAGGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGPTRLIPNLTYLGGQRDVNLKELI 180
Db 146 LAGGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGPTRLIPNLTYLGGQRDVNLKELI 205
QY 181 QONGIGYVLMASNTCPKDPFIPESHFLRVPVNDSPCEKILPWLDKSVDFIEKAKASNGCV 240
Db 206 QONGIGYVLMASNTCPKDPFIPESHFLRVPVNDSPCEKILPWLDKSVDFIEKAKASNGCV 265
QY 241 LVHCLAGISRSATITAIYIMKRMDSLDEAYRFVKEKPTTISPNFNLGQLDYEKKIKN 300
Db 266 LVHCLAGISRSATITAIYIMKRMDSLDEAYRFVKEKPTTISPNFNLGQLDYEKKIKN 325
QY 301 QT 302
Db 326 QT 327

RESULT 12
US-10-072-012-703
Sequence 703, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tcherenev, Vellizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernhusen, Bryan
APPLICANT: Patuturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Beha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072.012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PAM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 703
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-703

Query Match 99.2%; Score 1552; DB 12; Length 690;
Best Local Similarity 99.3%; Pred. No. 3.6e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
DB 26 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 85
QY 61 DKVLITELIQHSAKHKVIDCSQKVVVYDSSQDVASLSDCFLTVLLGLKESFNSVHL 120
DB 86 DKVLITELIQHSAKHKVIDCSQKVVVYDSSQDVASLSDCFLTVLLGLKESFNSVHL 145
QY 121 LAGGFAEFSRCFPGLCGKSTLVPCTISQCLPVANIGPTRLIPNLVYLGQDRVLNKL 180
DB 146 LAGGFAEFSRCFPGLCGKSTLVPCTISQCLPVANIGPTRLIPNLVYLGQDRVLNKL 205
QY 181 QONGIGVYNASTYCKPPDIPESHFLRVVNDSPCEKILPMIDKSVDFEKAASNGCV 240
DB 206 QONGIGVYNASTYCKPPDIPESHFLRVVNDSPCEKILPMIDKSVDFEKAASNGCV 265
QY 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIK 300
DB 266 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIK 325
QY 301 QT 302
DB 326 QT 327

RESULT 13
US-10-425-114-54204
Sequence 54204, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingtong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaka, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 54204
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB4119-028-H6_FLI.pep
US-10-425-114-54204

Query Match 99.2%; Score 1552; DB 12; Length 690;
Best Local Similarity 99.3%; Pred. No. 3.6e-147;

Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
DB 26 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 85
QY 61 DKVLITELIQHSAKHKVIDCSQKVVVYDSSQDVASLSDCFLTVLLGLKESFNSVHL 120
DB 86 DKVLITELIQHSAKHKVIDCSQKVVVYDSSQDVASLSDCFLTVLLGLKESFNSVHL 145
QY 121 LAGGFAEFSRCFPGLCGKSTLVPCTISQCLPVANIGPTRLIPNLVYLGQDRVLNKL 180
DB 146 LAGGFAEFSRCFPGLCGKSTLVPCTISQCLPVANIGPTRLIPNLVYLGQDRVLNKL 205
QY 181 QONGIGVYNASTYCKPPDIPESHFLRVVNDSPCEKILPMIDKSVDFEKAASNGCV 240
DB 206 QONGIGVYNASTYCKPPDIPESHFLRVVNDSPCEKILPMIDKSVDFEKAASNGCV 265
QY 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIK 300
DB 266 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIK 325
QY 301 QT 302
DB 326 QT 327

RESULT 14
US-10-072-012-681
Sequence 681, Application US/10072012
Publication No. US2004003493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Patuturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Radigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie E.
APPLICANT: Taupier, Jr, Raymond J.
APPLICANT: Gusev, Vladimr Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine B.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 681
LENGTH: 665
TYPE: PRF
ORGANISM: Homo sapiens
US-10-072-012-681

Query Match 98.8%; Score 1545; DB 12; Length 665;
Best Local Similarity 99.0%; Pred. No. 1.7e-146;
Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLSGETEKVLLIDSRPFVEYNTSHILEAININCKMKRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLSGETEKVLLIDSRPFVEYNTSHILEAININCKMKRRLOQ 60
QY 61 DKVLITELIQHSARKHKVDIDCSQKVVYDQSSQDVASLSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIQHSARKHKVDIDCSQKVVYDQSSQDVASLSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCGKSTLVPTCISQPCLPVANIPTLILPMLYLGGQRDVNLKELI 180
DB 121 LAGGFAEFSRCFPGLCGKSTLVPTCISQPCLPVANIPTLILPMLYLGGQRDVNLKELI 180
QY 181 QONGIGYVLNASYTCPEKDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLNASYTCPEKDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

RESULT 15

US-10-094-749-2312
Sequence 2312, Application US/10094749
Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUTKO
APPLICANT: HIO, YUKI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOKIKO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOYUKI
APPLICANT: NAGAHARA, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
PRIOR FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24

PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 2312
LENGTH: 665
TYPE: PRF
ORGANISM: Homo sapiens
US-10-094-749-2312

Query Match 98.8%; Score 1545; DB 15; Length 665;
Best Local Similarity 99.0%; Pred. No. 1.7e-146;
Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLSGETEKVLLIDSRPFVEYNTSHILEAININCKMKRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLSGETEKVLLIDSRPFVEYNTSHILEAININCKMKRRLOQ 60
QY 61 DKVLITELIQHSARKHKVDIDCSQKVVYDQSSQDVASLSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIQHSARKHKVDIDCSQKVVYDQSSQDVASLSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCGKSTLVPTCISQPCLPVANIPTLILPMLYLGGQRDVNLKELI 180
DB 121 LAGGFAEFSRCFPGLCGKSTLVPTCISQPCLPVANIPTLILPMLYLGGQRDVNLKELI 180
QY 181 QONGIGYVLNASYTCPEKDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLNASYTCPEKDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

Search completed: June 21, 2004, 13:24:58
Job time : 26.9214 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM proteain - protein search, using sw model

Run on: June 21, 2004, 13:20:38 ; Search time 10.3061 Seconds
(without alignments)
1512.797 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMIGTQIVTERLVALLE.....PNNFLGQLDYKRIKKT 302

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: Issued Patents AA:*
 - 2: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
 - 3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
 - 4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
 - 5: /cgn2_6/prodata/2/1aa/PTUS_COMB.pep:*
 - 6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1552	99.2	665	4	US-09-816-494-2
2	654.5	41.8	170	4	US-09-544-716-14
3	654.5	41.8	170	4	US-09-557-921-15
4	654.5	41.8	170	4	US-09-564-357-17
5	654.5	41.8	170	4	US-09-619-380-16
6	469	30.0	482	4	US-09-557-921-2
7	425	27.2	314	4	US-09-371-671B-11
8	425	27.2	367	2	US-08-990-379-6
9	425	27.2	384	2	US-08-530-290-23
10	425	27.2	394	4	US-09-702-705-805
11	425	27.2	394	4	US-09-702-705-805
12	425	27.2	394	4	US-09-736-457-805
13	425	27.2	394	4	US-09-736-457-827
14	425	27.2	394	4	US-09-614-124B-805
15	425	27.2	394	4	US-09-614-124B-805
16	425	27.2	394	4	US-09-611-325-805
17	425	27.2	394	4	US-09-671-325-805
18	425	27.2	394	4	US-09-589-184-805
19	425	27.2	394	4	US-09-589-184-827
20	423.5	27.1	395	2	US-08-990-379-5
21	415	26.5	367	2	US-08-530-290-24
22	409.5	26.2	313	2	US-08-990-379-7
23	409	26.2	314	3	US-09-164-193-22
24	409	26.2	314	4	US-09-221-448A-22
25	377.5	24.1	393	2	US-08-990-379-4
26	357	22.8	302	4	US-09-702-705-806
27	357	22.8	302	4	US-09-736-457-806

28	357	22.8	302	4	US-09-614-124B-806	Sequence 806, App
29	357	22.8	302	4	US-09-671-325-806	Sequence 806, App
30	357	22.8	302	4	US-09-589-184-806	Sequence 806, App
31	349.5	22.3	397	2	US-08-990-379-8	Sequence 8, Appl1
32	346.5	22.2	168	4	US-09-544-716-13	Sequence 13, Appl1
33	346.5	22.2	168	4	US-09-557-921-13	Sequence 13, Appl1
34	346.5	22.2	168	4	US-09-564-357-16	Sequence 16, Appl1
35	346.5	22.2	168	4	US-09-619-380-15	Sequence 15, Appl1
36	338.5	21.6	170	4	US-09-544-716-12	Sequence 12, Appl1
37	338.5	21.6	170	4	US-09-557-921-12	Sequence 12, Appl1
38	338.5	21.6	170	4	US-09-564-357-15	Sequence 15, Appl1
39	338.5	21.6	170	4	US-09-619-380-14	Sequence 14, Appl1
40	319	20.4	169	4	US-09-544-716-16	Sequence 16, Appl1
41	319	20.4	169	4	US-09-557-921-17	Sequence 17, Appl1
42	319	20.4	169	4	US-09-564-357-19	Sequence 19, Appl1
43	319	20.4	169	4	US-09-619-380-18	Sequence 18, Appl1
44	312	19.9	168	4	US-09-544-716-15	Sequence 15, Appl1
45	312	19.9	168	4	US-09-557-921-16	Sequence 16, Appl1

ALIGNMENTS

RESULT 1
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR APPLICATION NUMBER: 2001-03-23
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-494-2

Query Match	Score	DB	Length	Best Local Similarity	Pred. No.	Mismatches	Indels	Gaps
Query Match	99.2%	Score 1552	DB 4	Length 665	Best Local Similarity 99.3%	Pred. No. 3e-161	1	Indels 0
Matches 300	Conservative	1	Mismatches	0	Gaps	0		
QY	1	MAHEMIGTQIVTERLVALLESGTEKVLIDSRPFVEYNTSHLEATININCKLMKRLQ	60					
DB	1	MAHEMIGTQIVTERLVALLESGTEKVLIDSRPFVEYNTSHLEATININCKLMKRLQ	60					
QY	61	DKVLITELIHSKHKVDIDSGKVVVYDSSQDVASLSSDCLFVLLGLKESFNSVHL	120					
DB	61	DKVLITELIHSKHKVDIDSGKVVVYDSSQDVASLSSDCLFVLLGLKESFNSVHL	120					
QY	121	LAGFAEFRCPCGCLCEGKSTVPTISQPCLPVANIIGPRRIIPNLYLGCQRYVANKELI	180					
DB	121	LAGFAEFRCPCGCLCEGKSTVPTISQPCLPVANIIGPRRIIPNLYLGCQRYVANKELI	180					
QY	181	QONGIGVYVNASYCKPDPFIPESHFLRVVNDSPCEKILPMWDSKVDIFEKASNGCV	240					
DB	181	QONGIGVYVNASYCKPDPFIPESHFLRVVNDSPCEKILPMWDSKVDIFEKASNGCV	240					
QY	241	LVHCLAGISRSATIAIAYIMKMDMSLDEAYRVEKEKRTISNENFLQGLDYKRIKX	300					
DB	241	LVHCLAGISRSATIAIAYIMKMDMSLDEAYRVEKEKRTISNENFLQGLDYKRIKX	300					
QY	301	QT 302						
DB	301	QT 302						

RESULT 2
US-09-544-716-14
; Sequence 14, Application US/09544716
; Patent No. 6492157
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-9 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.415
; CURRENT APPLICATION NUMBER: US/09/544.716
; CURRENT FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 170
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-544-716-14

Query Match 41.8%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 1e-63;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;
QY 134 GLCEGK-STLVPTCISQPCLPVANIGPTRLIPMLYIGQCRDVLNKLIOQNGIGVYLNAS 192
DB 1 GLCEGKPAALLPMSLSQPCLPVPSVGLTRILPHLYIGSQDVLNKLDMQNGISYVLNAS 60
QY 193 YTCRKPDPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSA 252
DB 61 NSCPKDPFICESRMRVPINDNCEKLLPMLDKSIFETIDAKLSSQCVLVHCLAGISRSA 120
QY 253 TIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
DB 121 TIAIAYIMKTMGMSDDAYRFVDRRPSISPNNFLGQLLEVERTLK 167

RESULT 3
US-09-557-921-15
; Sequence 15, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT APPLICATION NUMBER: US/09/557.921
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-557-921-15

Query Match 41.8%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 1e-63;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;
QY 134 GLCEGK-STLVPTCISQPCLPVANIGPTRLIPMLYIGQCRDVLNKLIOQNGIGVYLNAS 192
DB 1 GLCEGKPAALLPMSLSQPCLPVPSVGLTRILPHLYIGSQDVLNKLDMQNGISYVLNAS 60
QY 193 YTCRKPDPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSA 252
DB 61 NSCPKDPFICESRMRVPINDNCEKLLPMLDKSIFETIDAKLSSQCVLVHCLAGISRSA 120
QY 253 TIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
DB 121 TIAIAYIMKTMGMSDDAYRFVDRRPSISPNNFLGQLLEVERTLK 167

RESULT 4

US-09-564-357-17
; Sequence 17, Application US/09564357
; Patent No. 6645753
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-5 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.413
; CURRENT APPLICATION NUMBER: US/09/564.357
; CURRENT FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 170
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-564-357-17

Query Match 41.8%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 1e-63;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;
QY 134 GLCEGK-STLVPTCISQPCLPVANIGPTRLIPMLYIGQCRDVLNKLIOQNGIGVYLNAS 192
DB 1 GLCEGKPAALLPMSLSQPCLPVPSVGLTRILPHLYIGSQDVLNKLDMQNGISYVLNAS 60
QY 193 YTCRKPDPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSA 252
DB 61 NSCPKDPFICESRMRVPINDNCEKLLPMLDKSIFETIDAKLSSQCVLVHCLAGISRSA 120
QY 253 TIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
DB 121 TIAIAYIMKTMGMSDDAYRFVDRRPSISPNNFLGQLLEVERTLK 167

RESULT 5
US-09-619-380-16
; Sequence 16, Application US/09619380
; Patent No. 6649391
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-11 DUAL SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.418
; CURRENT APPLICATION NUMBER: US/09/619.380
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 170
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-619-380-16

Query Match 41.8%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 1e-63;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;
QY 134 GLCEGK-STLVPTCISQPCLPVANIGPTRLIPMLYIGQCRDVLNKLIOQNGIGVYLNAS 192
DB 1 GLCEGKPAALLPMSLSQPCLPVPSVGLTRILPHLYIGSQDVLNKLDMQNGISYVLNAS 60
QY 193 YTCRKPDPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSA 252
DB 61 NSCPKDPFICESRMRVPINDNCEKLLPMLDKSIFETIDAKLSSQCVLVHCLAGISRSA 120
QY 253 TIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
DB 121 TIAIAYIMKTMGMSDDAYRFVDRRPSISPNNFLGQLLEVERTLK 167

RESULT 6
US-09-557-921-2

	Query Match	27.2%	Score	425;	DB	4;	Length	314;
	Best Local Similarity	34.3%	Pred No.	3,	7e-38;			
	Matches	104;	Conservative	48;	Mismatches	129;	Indels	22; Gaps
Oy								
D6	15 LVALISGTEKVLIDSRPFVEYNTSHIIEAININCSKLMMRLOODKYLTITLIQHSAK	74						
		:						
	13 LRLLIREGAOQLDDCRSFFAENFGHLAGSVVFSTIVRRAGANGLEIHVNAEIR	72						
		:						

Query Match	27.2%	Score 425;	DB 2;	Length 367;
Best Local Similarity	34.3%	Pred. No. 4, 6e-38;		
Matches	104;	Conservative	48;	Mismatches 129; Indels 22; Gaps 5

QY	15	LVALLESTGEKVLIDSRPFVEYNTSHILEAININCSKIMKRLODKVLTTELQHSAK	74
DB	13	LRALLREGAQCCLLDRCSPFAFNAGHLAGSVNVPFSTIVRRRAGAMGLEHTVPNAELR	72
QY	75	HKVDIDCSQKVVVVDQSDVASLSSDCEFLVTLGKL--EKSPFNVHLLAGGAERSRCF	132
DB	73	GRLLAGAHVAHVLLIDERSASLDGARKDGTLLAALACREARSTQYFFLOGGYEATSSASC	132
QY	133	PGLEGGKSTLVPTCISOP-----CLPVAANI-----GPTRIELNLYAGCCORDVL	175
DB	133	PELCSKOST--PTGISLPELSTSVPRPSASGSCSCSTPLVDGQPVETLSFLYUGSAVYHNS	190
QY	176	NKELLIQNGIGVYLNASYTCRKPDEFPSHFPLRVPNVDSFCSEKILPMWLKDSVDFIEKAYA	235
DB	191	RKQMLDALGITALINVANSCPN--HREGHYQXKSIPEVDHKKADISSWFBEALDFDISIDK	249
QY	236	SNNGCYLVHCLAGISRSATILALAYIMKRDMSLDEKAYRFPYKEKGRPTISSPNPNLGGQLLDE	295
DB	250	AGGRFVHCQAGISRSATILCLAYLMRTNRVKLKDEAFVYKORRSIISPNFSFWGQLQFE	309
QY	296	KKI	298
DB	310	SQV	312

RESULT 9

US-08-530-290-23
Sequence 23, Application US/08530290
Patent No. 5958721
GENERAL INFORMATION:
APPLICANT: Marshall, Christopher John
APPLICANT: Ashworth, Alan
APPLICANT: Hughes, David Anthony
TITLE OF INVENTION: Methods for Screening of Substances for
NUMBER OF SEQUENCES: 24
THERAPEUTIC ACTIVITY AND YEAST FOR USE THEREIN
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,290
FILING DATE: 14-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB94/00694
FILING DATE: 31-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9402573.1
FILING DATE: 10-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9307250.2
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 084611-0000000US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0200
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-530-290-23

Query Match 27.2%; Score 425; DB 2; Length 394;
Best Local Similarity 33.1%; Pred. No. 5.2e-38;

Matches 99; Conservative 62; Mismatches 116; Indels 22; Gaps 8;

QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCSKLMRRLOQDVLITELI--QHSAXHK 76
DB 39 LPSG-GKCLLDRCPLAHSAAGYILGSVNRNCTIVRR-AGKVSLEQLPABEEVRR 96
QY 77 VDIDCSQKVYVVDOSQDVASLSDCFLTVLGLKESF--NSVHLIAGFAEFRCFPG 134
DB 97 LRSGLYAVIVVDERSPRASLRSDSTVSLVQALRRNARITDILKGGYERFSSEYPE 156
QY 135 LCEGKSTL-----VPTCISOP-----CLPVA-----NIGPTRIIPNLYGCGQDVANKEL 179
DB 157 FCSKTKALAIAPPVPSPATREPLDLDSCSGTPLHDOEGVEILPFLYIGSAVHAARRDM 216
QY 180 IOONGIGVYVNASYTCPKDPFIPESHFLRVVNDSPCEKILPWLDKSVDFIEKAKASNGC 239
DB 217 LDALGITLALINVSQDCPN-HFEGHYQYKCIPEVDNKHADISSFWMAIEYILDVVKDCRGR 275

QY 240 VLVHCLAGISRSATITAIYIMKRMDSIDEAYRFVKEKRPITSPNPNLQGLDYEKKI 298
DB 276 VLVHQAIGISRSATITCLAYIMMKRVRLEBAFEFYKQRSIIISNFSMQLQFESGV 334

RESULT 10

US-09-702-705-805
Sequence 805, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darlick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 805
LENGTH: 394
TYPE: PRT
ORGANISM: Homo sapiens
US-09-702-705-805

Query Match 27.2%; Score 425; DB 4; Length 394;
Best Local Similarity 33.1%; Pred. No. 5.2e-38;
Matches 99; Conservative 62; Mismatches 116; Indels 22; Gaps 8;

QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCSKLMRRLOQDVLITELI--QHSAXHK 76
DB 39 LPSG-GKCLLDRCPLAHSAAGYILGSVNRNCTIVRR-AGKVSLEQLPABEEVRR 96
QY 77 VDIDCSQKVYVVDOSQDVASLSDCFLTVLGLKESF--NSVHLIAGFAEFRCFPG 134
DB 97 LRSGLYAVIVVDERSPRASLRSDSTVSLVQALRRNARITDILKGGYERFSSEYPE 156
QY 135 LCEGKSTL-----VPTCISOP-----CLPVA-----NIGPTRIIPNLYGCGQDVANKEL 179
DB 157 FCSKTKALAIAPPVPSPATREPLDLDSCSGTPLHDOEGVEILPFLYIGSAVHAARRDM 216
QY 180 IOONGIGVYVNASYTCPKDPFIPESHFLRVVNDSPCEKILPWLDKSVDFIEKAKASNGC 239
DB 217 LDALGITLALINVSQDCPN-HFEGHYQYKCIPEVDNKHADISSFWMAIEYILDVVKDCRGR 275
QY 240 VLVHCLAGISRSATITAIYIMKRMDSIDEAYRFVKEKRPITSPNPNLQGLDYEKKI 298
DB 276 VLVHQAIGISRSATITCLAYIMMKRVRLEBAFEFYKQRSIIISNFSMQLQFESGV 334

RESULT 11

US-09-702-705-827
Sequence 827, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darlick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 827
LENGTH: 394
TYPE: PRF
ORGANISM: Homo sapiens
US-09-702-705-827

Query Match 27.2%; Score 425; DB 4; Length 394;
Best Local Similarity 33.4%; Pred. No. 5,2e-38;

Matches 100; Conservative 64; Mismatches 113; Indels 22; Gaps 9;

QY 19 LBSGTEKVLIDSRPVEYNTSHLEAININCSKMKRRLQODKVLITELI--QHSAXKH 76
DB 39 LPSG-GKCLLDRCRPLASHAGYILGSVNVRCNTIVRR-AKGSVSLQGLPABEEVRR 96
QY 77 VDDCSQKVVYVDOSSQDVASLSDCFLTVLGLKLEKSF--NSVHLAAGFAFSCFP 134
DB 97 LBSGLSAVIVYDERPRAESLREDSTVSLVQALRRNERTDICKKGYSFSSSEYR 156
QY 135 LCEGKSTL-----VPTCISQ-----C-LPVANI-GPTRILPNLYGQORDVINKEL 179
DB 157 FCSKTRALAAIPPPVPASATEPLDLCSSCGTFLHDGCGFVELLPFLYLSAVHAARRDM 216
QY 180 IQONGIGYVLNASYTCPEKDPFIPESHFLAVPNVDSCEKILPWLDSVDFIEKAKSNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQKCIPEVDNKHADISSWFMEIETIDAVKOCGR 275
QY 240 VLVHCLAGISRSATIAIAYIMKMDMSLDEAVRFVEKERTPTSPNPNFLGQLIDYEKI 298
DB 276 VLVHCQAGISRSATICTAIYIMMKRVRLAEAFEFVQORSIIISPNFSFGQLQFESSQV 334

RESULT 12

US-09-736-457-805
Sequence 805, Application US/09736457
Patent No. 6509448

GENERAL INFORMATION:

APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaityanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darlick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 805
LENGTH: 394
TYPE: PRF
ORGANISM: Homo sapiens
US-09-736-457-805

Query Match 27.2%; Score 425; DB 4; Length 394;
Best Local Similarity 33.4%; Pred. No. 5,2e-38;

Matches 99; Conservative 62; Mismatches 116; Indels 22; Gaps 8;

QY 19 LBSGTEKVLIDSRPVEYNTSHLEAININCSKMKRRLQODKVLITELI--QHSAXKH 76
DB 39 LPSG-GKCLLDRCRPLASHAGYILGSVNVRCNTIVRR-AKGSVSLQGLPABEEVRR 96

QY 77 VDDCSQKVVYVDOSSQDVASLSDCFLTVLGLKLEKSF--NSVHLAAGFAFSCFP 134
DB 97 LBSGLSAVIVYDERPRAESLREDSTVSLVQALRRNERTDICKKGYSFSSSEYR 156
QY 135 LCEGKSTL-----VPTCISQ-----C-LPVANI-GPTRILPNLYGQORDVINKEL 179
DB 157 FCSKTRALAAIPPPVPASATEPLDLCSSCGTFLHDGCGFVELLPFLYLSAVHAARRDM 216
QY 180 IQONGIGYVLNASYTCPEKDPFIPESHFLAVPNVDSCEKILPWLDSVDFIEKAKSNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQKCIPEVDNKHADISSWFMEIETIDAVKOCGR 275
QY 240 VLVHCLAGISRSATIAIAYIMKMDMSLDEAVRFVEKERTPTSPNPNFLGQLIDYEKI 298
DB 276 VLVHCQAGISRSATICTAIYIMMKRVRLAEAFEFVQORSIIISPNFSFGQLQFESSQV 334

RESULT 13

US-09-736-457-827
Sequence 827, Application US/09736457
Patent No. 6509448

GENERAL INFORMATION:

APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaityanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darlick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 827
LENGTH: 394
TYPE: PRF
ORGANISM: Homo sapiens
US-09-736-457-827

Query Match 27.2%; Score 425; DB 4; Length 394;
Best Local Similarity 33.4%; Pred. No. 5,2e-38;

Matches 100; Conservative 64; Mismatches 113; Indels 22; Gaps 9;

QY 19 LBSGTEKVLIDSRPVEYNTSHLEAININCSKMKRRLQODKVLITELI--QHSAXKH 76
DB 39 LPSG-GKCLLDRCRPLASHAGYILGSVNVRCNTIVRR-AKGSVSLQGLPABEEVRR 96
QY 77 VDDCSQKVVYVDOSSQDVASLSDCFLTVLGLKLEKSF--NSVHLAAGFAFSCFP 134
DB 97 LBSGLSAVIVYDERPRAESLREDSTVSLVQALRRNERTDICKKGYSFSSSEYR 156
QY 135 LCEGKSTL-----VPTCISQ-----C-LPVANI-GPTRILPNLYGQORDVINKEL 179
DB 157 FCSKTRALAAIPPPVPASATEPLDLCSSCGTFLHDGCGFVELLPFLYLSAVHAARRDM 216
QY 180 IQONGIGYVLNASYTCPEKDPFIPESHFLAVPNVDSCEKILPWLDSVDFIEKAKSNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQKCIPEVDNKHADISSWFMEIETIDAVKOCGR 275
QY 240 VLVHCLAGISRSATIAIAYIMKMDMSLDEAVRFVEKERTPTSPNPNFLGQLIDYEKI 298
DB 276 VLVHCQAGISRSATICTAIYIMMKRVRLAEAFEFVQORSIIISPNFSFGQLQFESSQV 334

RESULT 14

US-09-614-124B-805
Sequence 805, Application US/09614124B

```
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B--805
```

```
Query Match          27.2%; Score 425; DB 4; Length 394;
Best Local Similarity 33.1%; Pred. No. 5,2e-38;
Matches 99; Conservative 62; Mismatches 116; Indels 22; Gaps 8;
```

```
QY 19 LESGTEKVLIDSRPFVEYNTSHILEAININCSKIMKRRLOQDKVLTIELI--QHSAGHK 76
DB 39 LPSSG-GKCLLDRCRFLHAGYILGSVNVRCNTIVRRR-AKGSVSLQIILPAEEVVRAR 96
QY 77 VIDDSQKVVYVDSSQDVASISDCFLTVILGKLEKSF--NSVHLLAGFAEFSCRCPG 134
DB 97 LRSGLYSAVIYVDERSPRAESLRSDSTVSLVQALRNAERTDICTLKGGERFSSEYPE 156
QY 135 LQEGKSTL-----VPTCISOP-----CLPVA-----NIGPTIILPNLYIGCCRDVANKEL 179
DB 157 FCSKTKALAIAPPVPSPATEPLDLDGSCGTFPLHDGCPVRIILPFLYLGSAVYHARRDM 216
QY 180 IOONGIGVYLAASYCPKPDFTPESHFLRVPNVDSFCEKILPWLKSDVFIEKAKASNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQYKCIPEVDNHRKADISSWEMEALEYIDAVIDCRGR 275
QY 240 VLVHCLAGISRSATTAIAYIMKRMDSIDEAYRFYKRRPTISPNNFLGQLLDYEKKI 298
DB 276 VLVHCQAGISRSATICTLAYIMKKGRVLEAEFVYKRRSTISPNSFWGQLLOFESQV 334
```

RESULT 15

```
US-09-614-124B-827
; Sequence 827, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 827
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B-827
```

```
Query Match          27.2%; Score 425; DB 4; Length 394;
Best Local Similarity 33.4%; Pred. No. 5,2e-38;
Matches 100; Conservative 64; Mismatches 113; Indels 22; Gaps 9;
```

```
QY 19 LESGTEKVLIDSRPFVEYNTSHILEAININCSKIMKRRLOQDKVLTIELI--QHSAGHK 76
DB 39 LPSSG-GKCLLDRCRFLHAGYILGSVNVRCNTIVRRR-AKGSVSLQIILPAEEVVRAR 96
QY 77 VIDDSQKVVYVDSSQDVASISDCFLTVILGKLEKSF--NSVHLLAGFAEFSCRCPG 134
DB 97 LRSGLYSAVIYVDERSPRAESLRSDSTVSLVQALRNAERTDICTLKGGERFSSEYPE 156
QY 135 LQEGKSTL-----VPTCISOP-----CLPVA-----NIGPTIILPNLYIGCCRDVANKEL 179
DB 157 FCSKTKALAIAPPVPSPATEPLDLDGSCGTFPLHDGCPVRIILPFLYLGSAVYHARRDM 216
QY 180 IOONGIGVYLAASYCPKPDFTPESHFLRVPNVDSFCEKILPWLKSDVFIEKAKASNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQYKCIPEVDNHRKADISSWEMEALEYIDAVIDCRGR 275
QY 240 VLVHCLAGISRSATTAIAYIMKRMDSIDEAYRFYKRRPTISPNNFLGQLLDYEKKI 298
DB 276 VLVHCQAGISRSATICTLAYIMKKGRVLEAEFVYKRRSTISPNSFWGQLLOFESQV 334
```

```
Search completed: June 21, 2004, 13:25:44
Job time : 11.3061 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM proteoin - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 21:02:42 ; Search time 3544.67 Seconds
(without alignments)
3692.750 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMGTGIVTERLVALL.....PNNFUGOLDYEKIKNOT 302

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xjh
-Q/cg2.1/USPTO.spool/US10029345/runcat.21062004.164146.29030/app.query.fasta.1.1294
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LCOFCL=0 -LCOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=pro -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345.@CG2.1.1.6617@runcat.21062004.164146.29030 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_ov:.*
5: gb_ov:.*
6: gb_ov:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_stb:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*
15: em_ba:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_mu:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pac:.*
24: em_ph:.*
25: em_pl:.*
26: em_ro:.*
27: em_stb:.*
28: em_un:.*

29: em_vl:.*
30: em_hcg_hum:.*
31: em_hcg_inv:.*
32: em_hcg_other:.*
33: em_hcg_mus:.*
34: em_hcg_pin:.*
35: em_hcg_rod:.*
36: em_hcg_mam:.*
37: em_hcg_vrt:.*
38: em_sy:.*
39: em_hcgo_hum:.*
40: em_hcgo_mus:.*
41: em_hcgo_other:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1564	100.0	5450	6 AX482439	AX482439 Sequence
2	1564	100.0	5450	6 AX482478	AX482478 Sequence
3	1552	99.2	1998	6 AX260342	AX260342 Sequence
4	1552	99.2	2071	6 AX921917	AX921917 Sequence
5	1552	99.2	2118	6 AX099933	AX099933 Sequence
6	1552	99.2	2732	6 AX180875	AX180875 Sequence
7	1552	99.2	3059	6 AX278461	AX278461 Sequence
8	1552	99.2	3104	6 AX405700	AX405700 Sequence
9	1552	99.2	3284	6 BC042101	BC042101 Homo sapi
10	1552	99.2	3496	6 AX441210	AX441210 Sequence
11	1552	99.2	3521	9 AB052156	AB052156 Homo sapi
12	1552	99.2	3544	6 AX260340	AX260340 Sequence
13	1552	99.2	3566	6 AF506796	AF506796 Homo sapi
14	1552	99.2	3766	6 AX374994	AX374994 Sequence
15	1552	99.2	4790	6 BD171157	BD171157 Novel gen
16	1552	99.2	4790	6 BD183432	BD183432 Novel gen
17	1552	99.2	4790	9 AB051487	AB051487 Homo sapi
18	1545	98.8	2102	6 AX713989	AX713989 Sequence
19	1545	98.8	2102	6 AK055973	AK055973 Homo sapi
20	1544	98.7	5111	6 AX482372	AX482372 Sequence
21	1517	97.0	4874	10 AB052157	AB052157 Mus muscu
22	1517	97.0	4943	10 BC059232	BC059232 Mus muscu
23	1517	97.0	4975	10 BC057321	BC057321 Mus muscu
24	1514	96.8	4992	10 AF345951	AF345951 Mus muscu
25	1504.5	96.2	2200	6 AX921915	AX921915 Sequence
26	1492	95.4	2756	6 AX482444	AX482444 Sequence
27	1375.5	87.9	4026	10 AF345953	AF345953 Mus muscu
28	1312	83.9	2807	9 BC031643	BC031643 Homo sapi
29	1246	79.7	749	6 AX868469	AX868469 Sequence
30	1246	79.7	749	6 BD148531	BD148531 Primer fo
31	1194.5	76.4	1935	9 AY038927	AY038927 Homo sapi
32	1194.5	76.4	3332	6 AX441229	AX441229 Sequence
33	1172.5	75.0	4827	10 AF345952	AF345952 Mus muscu
34	1034	66.1	3861	10 AF345954	AF345954 Mus muscu
35	1019.5	65.2	3231	5 BC063941	BC063941 Dantio rer
36	989	63.2	2264	9 BC045110	BC045110 Homo sapi
37	988	63.2	2377	6 AR255944	AR255944 Sequence
38	988	63.2	2377	9 HS027193	HS027193 Human prote
39	985	63.0	2453	6 A59888	A59888 Sequence 6
40	985	63.0	2453	10 MANTT11GN	X95518 M.musculu
41	985	63.0	2600	10 BC052705	BC052705 Mus muscu
42	943	60.3	2476	6 AX552119	AX552119 Sequence
43	934.5	59.8	2415	6 A59887	A59887 Sequence 5
44	858.5	54.9	183797	9 AC022400	AC022400 Homo sapi
45	858.5	54.9	198218	2 AC037447	AC037447 Homo sapi

RESULT 1

ALIGNMENTS

AX482439
LOCUS AX482439 5450 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 108 from Patent WO02057460.
ACCESSION AX482439
VERSION AX482439.1 GI:22316984
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS
1 Todderud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G.,
Suchard,S., Banas,D., Bassolino,D., Feder,J., Krystek,S.,
Mcatee,P., Mintier,G., Siemers,N., Jackson,D.G. and Ramamathan,C.
TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 108 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)
FEATURES
Location/Qualifiers
1..5450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538..2535
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD44451.1"
/db_xref="GI:22316985"
/translation="MAHEMIGTQIVTERLVALLSGETRKVLLIDSRPFVEYNTSHILE
ATINCSKLMKRRLOODKVLTELIOHSHKHVDIDCQKVVVDQSDVASISDC
FTVLVLGKESFNSVHLIAGFAFSSFCFGLCEGKSTLPTCISQPLVANI
RILPNLYLGCCQDQVINKELIQONGIGYVLAISYTCPKDPIDESHFLVNPDSFCEK
ILPMDKSDVPLEKASNGCVLVHCLAGISRSATIAIAYIMKMDMSLDNAYFVEK
KPTISPNPFLGOLLDEYKIKNOTGASGPKSKILHLEKPMPPVAYSGCKSR
TLPSPCADSATSEAGORPVHPASVSPVPSVPSBILDSVLVALSGLIHSADULEB
SNKRSRSLDKISYSASMAASLHAGSSSDMLERYKBSITTDGNTKLCQFSPVGE
LSEQPTSPSPDEKASITPKUQTAKPSDSQSKRLHSVTSSSGAGSLSPLRSGS
VEDNHTSFLPGLSTSQOHLTKSNAGLGYKMHSDILAPOSTPBLTSSWYFATSSHF
YSASAIYGGSAISYASVSCOLPTCGDQYVSRROKPEDRADSRSSWHEBSPPFKQK
RRSCOMERGEISIMENSRRELGKVGSSPSGSEIIEVS"

ORIGIN
Alignment Scores:
Pred. No.: 6,58e-151 Length: 5450
Score: 1564.00 Matches: 302
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x AX482439 (1-5450)

QY 1 MetAlAHISGIUMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 538 ATGGCCCATGAGCATTTGGAACCTCAATTCTTACTGAGAGGTGGGCGCTCTGCTGAA 597
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluIleAsnThrSer 40
DB 598 AGTGAACGGAAGAGTCTCTAATTGATGCGCGGCACTTGTGGAAATCAATACATCC 657
QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIleSerIleMetLeuValArgLeuGln 60
DB 658 CACATTTGGAGGCCATTAAATATCAACTGCTCCCAAGCTTATGAGCGAAGGTTCACACG 717
QY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
DB 718 GACAAAGGTTAATTACAGGCTCATCCAGCATTCACCGAAACATAAGTTGACATTGAT 777
QY 81 CysSerGlnIleValIleValIleAspGlnSerSerGlnAspValAlaSerIleAsnSer 100
DB 778 TGCAGTCAGAGGTTGATTGATTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 837
QY 101 AspCysPheLeuThrValLeuLeuGluIleValSerPheAsnSerValHisIleu 120

DB 838 GACTGTTTTCACCTGATCTTCTGGTAAACTGGAGAAAGACTTCAACTGTTCACTTG 897
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGluGlyIleSer 140
DB 898 CTTCAGAGTGGGTTTGGCTGAGTTCTTCGTTGTTTCCCTGGGCTCTGTGAAAGAAATCC 957
QY 141 ThrIleValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 958 ACTCTAGTCCCTACCTCACTTCTCAGCTTGCTTACTGTTGCAACATTTGGGCAACC 1017
QY 161 ArgIleLeuProAsnLeuIleGlyIleValLeuAsnAlaSerIleThrCysProIleAspPhe 200
DB 1078 CAGCAGATGGAGTTGGTTATGCTTAAATGCCAGCTATCTGTCGAAAGCTTGACTTT 1137
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
DB 1138 ATCCCGAGTCTCATTTCTGCGTGCCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1197
QY 221 ProThrLeuAspIleSerValAspPheIleGluValAlaValAlaSerIleGlyVal 240
DB 1198 CCGTGGTTGACAAATCATGATTCATTGAGAAAGCAAAAGCTTCAATGATGTGTT 1257
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleVal 260
DB 1258 CTAGTGACATGTTTACGTGGATCTCCCGCTCGGCCACCATGCTTATGCTTACATCAG 1317
QY 261 LysArgMetAspMetSerLeuAspGluAlaIleIleValPheValIleGluIleAspProThr 280
DB 1318 AAGAGATGACATGCTTATGATGAGCTTACAGATTTGTGAAAGAAAAAGACTTACT 1377
QY 281 IleSerProAsnPheAsnPheLeuGluIleGluLeuAspIleGluIleValIleAsn 300
DB 1378 ATATCTCCAAACTTCAATTTTCTGGCCCAACTCTGACTATGAGAAAGATTAGAAGC 1437
QY 301 GlnThr 302
DB 1438 CAGACT 1443
RESULT 2
AX482478
LOCUS AX482478 5450 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 147 from Patent WO02057460.
ACCESSION AX482478
VERSION AX482478.1 GI:22316999
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS
1 Todderud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G.,
Suchard,S., Banas,D., Bassolino,D., Feder,J., Krystek,S.,
Mcatee,P., Mintier,G., Siemers,N., Jackson,D.G. and Ramamathan,C.
TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 147 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)
FEATURES
Location/Qualifiers
1..5450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538..2535
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD44453.1"
/db_xref="GI:22317000"
/translation="MAHEMIGTQIVTERLVALLSGETRKVLLIDSRPFVEYNTSHILE

ATINSGKIMKRLRLOODKVLITELIOHSAKHKVIDCSOKVVVVYDOSODVASTSDC
 FLTVLGLKEKSPNSHYHLAAGFAERSRCPGLCEKSTLVPICISQPCPLPANTICPT
 RILPNLYLGGQDVLNKEKLIQQNGIGVLAASATTCPEKPIPSHRLPVAIVDSFCEK
 ILPMIDKSVDFLEKAKASNGCVLHCLAGISRSATLAIYIMKRMASIDEAARFKE
 KRPTISPNFNFILQOLIDYEKIKNOTGASPKSKLKLHLKXENBVSIVSEGGQSE
 TPLSPCADSATSEAGORPVHPASVPSVOPSLLEDSPLVOALSGHLSDRLD
 SNLTKRSPSLDIKISVSASMAHSGFSSERDLHYVPSRTLLDCTNKLCOSSPQVE
 LSEQTPETSPDKERASIPKKLQOTARPSPDSORSLHVSRTSSCTPARRSLSPHRS
 VEDNHYTSPLFGISTSQOHLTKTSAGLGLKMSHSDIAPQSTPLSTSMYFATSEHG
 YSASAIYGGASAYSACSQPLPCQDQVYSVRROKPSRADPSRMSHEESPKEQPK
 RRSQMEFBSIWSNRSREELGKVGSGSFSGSMELIEVS"

ORIGIN

Alignment Scores:

Pred. No.:	6.58e-151	Length:	5450
Score:	1564.00	Matches:	302
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345A-109_COPY_1_302 (1-302) x AX482478 (1-5450)

QY 1 MetAlaHigLumetiLeglyThrgInileValThrgLuargleuValAlaleuLeuGlu 20
 DB 538 ATGGCCCATGAGATGATGAGAACTCAATTTGTTACTGAGAGTTGGCTCTGCTGGAA 537
 QY 21 SerGlyThrgLuarValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 DB 598 AGTGAAGAGGAAAGAGTCTGCTAATGATGAGCCGCTATTTGGAAATCAATACATCC 657
 QY 41 HisIleLeuGluAlaleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 DB 658 CACATTTTGAACACCATTAATCACTGCTCCAGCTTATGAGAGCAAGCTTGCAACAG 717
 QY 61 AspLysValLeuLeuLeuThrgLuarLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
 DB 718 GACAAAGTGTAAATTACAGAGCTCATTCACGAAACATTAAGGTTGACATTGAT 777
 QY 81 CysSerGlnLysValValValValValValValValValValValValValValValVal 100
 DB 778 TGCAGTCAGAGGTTGATTTAGATCAAGCTCCCAAGATGTTGGCTCTCTCTTCA 837
 QY 101 AspCysPheLeuThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
 DB 838 GACTGTTTCTCACTGACTTCTGGGTAACCTGAGAGAGCTTCACTGTTCACTG 897
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuLeuLeuLeuLeuLeu 140
 DB 898 CTTCGAGAGTGGTTGCTGAGTCTCTGTTGTTTCCCTGAGCTCTGTGAGAGAAATCC 957
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 958 ACTCTAGTCCCTACCTGCTCAGCTTGGCTTACCTGTGCAACATTTGGGCCCAAC 1017
 QY 161 ArgIleLeuProAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
 DB 1018 CGAATCTTCCCAATCTTTATCTTGGCTGCAAGGAGATGCTCTCAACAGAGCTGATA 1077
 QY 181 GluGlnAsnGlyIleGlyValLeuAsnAlaSerValThrCysProLysProAspPhe 200
 DB 1078 CAGCAGATGGAGTGTATGTTTAAATGCAAGCTTAACCTGTCCAAAGCTTGACTTT 1137
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleLeu 220
 DB 1138 ATCCCGAGTCTATTTCTCGCGGTGCTGTAATGACAGCTTTTGGAGAAATTTTG 1197
 QY 221 ProTTrpLeuAspLysSerValAspPheIleGluValAlaLysAsnGlyCysVal 240
 DB 1198 CCGGTGTGACAAATCAGATGATTTCTTGAAGAAAGCAAGCTTCAATGATGTT 1257
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaValIleMet 260

DB 1258 CTAGTCAGCTGTTAGCTGGATCTCCGCTCCGACCATGCTATCCCTACATCATG 1317
 QY 261 LysAlaGlyMetAspMetSerLeuAspGluAlaValThrgPheValIleGluLysAspProThr 280
 DB 1318 AAGAGATGAGACATGCTTTAATGATGAGGTTACAGATTTGTGAAGAAAAAGACTTACT 1377
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyGluLysIleLeuAsn 300
 DB 1378 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGACTATGAGAGAAAGATTAAAGAAC 1437
 QY 301 GlnThr 302
 DB 1438 CAGACT 1443

RESULT 3
 AX260342
 LOCUS AX260342 1998 bp DNA linear PAT 26-OCT-2001
 DEFINITION Sequence 3 from Patent WO0173059.
 ACCESSION AX260342
 VERSION AX260342.1 GI:16509305
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Meyers, R.A.
 38692 and 21117: dual specificity phosphatase molecules and uses
 therefor
 Patent: WO 0173059-A 3 04-OCT-2001;
 Millennium Pharmaceuticals, Inc. (US)

FEATURES

source location/Qualifiers
 1..1998
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 3.19e-150 Length: 1998
 Score: 1562.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 1
 Query Match: 99.23% Indels: 0
 DB: 6 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AX260342 (1-1998)

QY 1 MetAlaHigLumetiLeglyThrgInileValThrgLuargleuValAlaleuLeuGlu 20
 DB 1 ATGGCCCATGAGATGATGAGAACTCAATTTGTTACTGAGAGTTGGCTCTGCTGGAA 60
 QY 21 SerGlyThrgLuarValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 DB 61 AGTGAAGAGGAAAGAGTCTGCTAATGATGAGCCGCTATTTGGAAATCAATACATCC 120
 QY 41 HisIleLeuGluAlaleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 DB 121 CACATTTTGAAGCCCTTAATATCACTCTCCAACTTAATGAGAGAGGTTGCAACAG 180
 QY 61 AspLysValLeuLeuThrgLuarLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
 DB 181 GACAAAGTGTAAATTACAGAGCTCATTCAGAGCAATTAAGGTTGACATTGAT 240
 QY 81 CysSerGlnLysValValValValValValValValValValValValValValValVal 100
 DB 241 TGCAGTCAGAGGTTGATTTAGATCAAGCTCCCAAGATGTTGGCTCTCTCTTCA 300
 QY 101 AspCysPheLeuThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
 DB 301 GACTGTTTCTCACTGACTTCTGGGTAACCTGAGAGAGCTTCACTGTTCACTG 360
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuLeuLeuLeuLeuLeu 140

|||||
Db 361 CTTCAGAGTGGGTTGGTGAAGTCTCTCTGTTTCCCTGGCTCTGTGAAGAAATCC 420
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 421 ACTGAGTCCCTACCTGATTCGATTCACGCTTGTACCTGTCGCAACATTGGGCCAAC 480
QY 161 ArgIleLeuProAsnLeuValCysGlnArgAspValLeuAsnGlyLeuIle 180
Db 481 CGAATTCCTCCCATCTTATCTTGGCTGCGAGAGATGCTCAACAAAGAGCTGATG 540
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnIleSerTyrThrCysProLysProAspPhe 200
Db 541 CAGCAGATGGGATGGTATGTTGTTAATGACGACGCAATACCTGTCAAAAGCTGACTT 600
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIleLeu 220
Db 601 ATCCCGAGTCTCATTTCTGCGGTGCTGTAATGACAGCTTTGTGAGAAATTTTG 660
QY 221 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
Db 661 CCGTGGTTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCAAATGATGTT 720
QY 241 LeuValHisCysLeuValIleGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db 721 CTAGTGACTGTTAGCTGGAGATCTCCGCTCGCACCATCGCTATCGCTACATCATG 780
QY 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlnLysArgProThr 280
Db 781 AAGAGATGGACATGCTTTAGATGAGCTTACAGTTTGTAAGAAAGAAAGAAAGCTTACT 840
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysIleLysAsn 300
Db 841 ATATCTCCAAACTTCATTTCTGGGCGCAACTCTGAGCTATGAGAAAGATTAAAGAC 900
QY 301 GlnThr 302
Db 901 CAGACT 906

RESULT 4
AX921917 2071 bp DNA linear PAT 18-DEC-2003
LOCUS AX921917
DEFINITION Sequence 257 from Patent WO02068649.
ACCESSION AX921917
VERSION AX921917.1 GI:40215410
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Patent: WO 02068649-A 257 06-SEP-2002;
JOURNAL Curagen Corporation (US)
FEATURES
source location/Qualifiers
1..2071
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.: 3,34e-150 Length: 2071
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservatave: 1
Best Local Similarity: 99.34% Mismatches: 0
Query Match: 99.23% Indels: 0
DB: 6 Gaps: 0

US-10-029-345a-109_copy_1_302 (1-302) x AX921917 (1-2071)
QY 1 MetValHisGlnMetIleGlyTyrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
|||||

Db 61 ATGCCCATGAGATGATGGAACTCAATTGTACTGAGAGTGTGTGCTCTGCGAA 120
QY 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlyTyrAsnThrSer 40
Db 121 AGTGAACGGAAGAAAGTCTCTCTTAATTGATAGCGGCGCAATTTGGAAATCAATACATCC 180
QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgTyrGlnGln 60
Db 181 CACATTTGGAGAGCATTAATATCACTGCTCCAGCTTATGAAAGCAAGAGTTGCAACAG 240
QY 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 241 GACAAAGTTTAATTAACAGAGCTCATTCACATTCACGCAAAATAGAGTTGACATTTGAT 300
QY 81 CysSerGlnLysValValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 301 TGCAATCAGAAAGTTGATAGTTAGTACCAAGCTCCCAAAATGTGCTCTCTCTTCA 360
QY 101 AspCysPheLeuThrValIleLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 361 GACTGTTTCTCACTGACTTCTGGGTAACTGGAAGAGGCTTCAACTCTGTTCACTG 420
QY 121 LeuAlaGlyLysPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
Db 421 CTTCGAGTGGGTTGTGAGTTCTCTGTTTTCCTGGCTCTGTGAAGAAATCC 480
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 481 ACTTACTCTCCCTACCTGCAATTTCTCAGCTTGTCTTACTGTTGCCAAATCTGGGCAACC 540
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGlnLeuIle 180
Db 541 CGAATCTTCCCAATCTTATCTTGTGCTGCGAGAGATGCTCTCAAAAGAGCTGATG 600
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnIleSerTyrThrCysProLysProAspPhe 200
Db 601 CAGCAGAAATGGAGTGGTATGTTAATGCAAGAAATACCTGTCCAAAGCTTACATT 660
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
Db 661 ATCCCGAGTCTCATTTCTGCGGTGCTGGAATGACAGCTTTGTGAGAAATTTTG 720
QY 221 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
Db 721 CCGTGGTTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCAAATGATGTT 780
QY 241 LeuValHisCysLeuValIleGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db 781 CTAGTGACTGTTTGTAGCTGGATCTCCGCTCGCACCATCGCTATCGCTACATCATG 840
QY 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlnLysArgProThr 280
Db 841 AAGAGATGGACATGCTTTAGATGAGCTTACAGATTGTGAAAGAAAGAAAGCTTACT 900
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysIleLysAsn 300
Db 901 ATATCTCCAAACTTCATTTCTGGGCGCAACTCTCGAGCTATGAGAAAGATTAAAGAC 960
QY 301 GlnThr 302
Db 961 CAGACT 966

RESULT 5
AX099933 2118 bp DNA linear PAT 02-APR-2001
LOCUS AX099933
DEFINITION Sequence 15 from Patent WO0120004.
ACCESSION AX099933
VERSION AX099933.1 GI:13538943
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Yue, H., Tang, Y. T., Bandman, O., Hillman, J. L., Baughn, M. R.,
Azimail, Y. and Lu, D. A.
TITLE Protein phosphatase and kinase proteins
JOURNAL Patent: WO 012004-A 15 22-MAR-2001;
Incyte Genomics, Inc. (US)
FEATURES
source 1. 2118
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID NO: 1234795CB1"
ORIGIN
Alignment Scores:
Pred. No.: 3,44e-150 Length: 2118
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 6 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x AX099933 (1-2118)
QY 1 MetAlAHieglumeti1egiYThrgInileValThrgluarphleuValAlaleuLeuglu 20
DB 25 ATGGCCCATGAGATGATTGGAACTCAATGTTACTGAGAGGTTGGCTCTGCTGGA 84
QY 21 SerGIYThrgluYsValleuLeuileaspSerArpProPhaValglurYraenThrSer 40
DB 85 AGTGAACCGGAAAGAGTGGCTGTAATGATGACCGGCACTTGTGGAAATCAATATC 144
QY 41 HieileuGlunAlaleanilleanilleanCySerlyLeuMetlySarTargleuIngIn 60
DB 145 CACATTTTGGAAACCAATTAATCACTGCTCAAGCTTAATGAAGCAAGTTGCAACG 204
QY 61 AspLyValleuileThrgluLeuileglnHieSerAlaYshIlyValAspIleap 80
DB 205 GACAAAGTGTAAATTACAGACTCAATCCACATTCACGCAACATAGGTTGACATTGAT 264
QY 81 CySerGlnlyeValValValValTyraSpGlnSerSerGlnAspValAlaSerLeuSer 100
DB 265 TGCAGTCAGAGGTTGATGATTACGATCAAGCTCCCAAGATGTTGCTCTCTTCA 324
QY 101 AspCyPheleuthrValleuLeuGlylyleuGluYsSerPheAsnSerValHleu 120
DB 325 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGCTTCAACTGTTCACTG 384
QY 121 LeuAlAGlylphelAGluPheserArgCyPheProglYleuCySgInuGlylySer 140
DB 385 CTTCAGAGTGGTTGCTGAGTCTCTCGTGTTCCTGCTGCTGTAAGGAAATCC 444
QY 141 ThleuValProThrCySileSerGlnProCyLeuProValAlaAnilleGlyProthr 160
DB 445 ACTCTAGTCCCTACCTGCACTTCTGAGCTTGTCTACTCTGTCACCAATGGGCCAAC 504
QY 161 ArgIleuProAsnleuYlyleuGlyCySglnArGspValleuAnlySgIleu 180
DB 505 CGAATCTTCCCAATCTTATCTTGGCTGCCAGAGATGCTCTCAACAAGAGAGCTGATG 564
QY 181 GluGlnAsnGlylleglyYrValleuAsnAlaSerTyThrCySpPolySProAspPhe 200
DB 565 CACGAGATGGAGTGGTATGTTAAATGCGAGCAATACCTGTCCAAAGCCCTGACTTT 624
QY 201 IleProGlnSerIshPheleuArgValProValAsnAspSerPheCySgIlySileu 220
DB 625 ATCCCGAGTCTATTTCTCGCGGTGCTCTGATAGACAGCTTTGTGGAAATTTTG 684
QY 221 ProTTrleuAspYsSerValAspPheileGluYsAlaYsAlaSerAnGlyCyVal 240
DB 685 CCGGTGGTGGACAAATCAATGATTTCAATGAGAAAGCAAAAGCTCCCAATGATGTGTT 744
QY 241 LeuValHieCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaYrIleMet 260

DB 745 CTAGTCACTGTTTACTGAGTCTCCGCTCCGACCATGCTATCGCTCATCATG 804
QY 261 LybArgMetAspMetSerleuAspGluAlaYrArgPheVallySgIlySargProthr 280
DB 805 AAGAGATGATGACATGCTTTAGATGAAGCTTACAGATTGTGMAAAGAAAAAAGCTACT 864
QY 281 IleSerProAsnPhesnPhelenuGlylSleuLeuAspTyGluYsYsIleYAsn 300
DB 865 ATATCTCCAAACTTCAATTTCTGGCCCAACTCTGACTATGAGAGAGATTAAAGAC 924
QY 301 GlnThr 302
DB 925 CAGACT 930
RESULT 6
AX180875 2732 bp DNA linear PAT 06-AUG-2001
LOCUS
DEFINITION Sequence 2 from Patent WO0146394.
ACCESSION AX180875
VERSION AX180875.1 GI:15132703
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Plozman, G. D., Martinez, R., Whyte, D., Manning, G., Sudarshanam, S.,
Hill, R. J. and Flanagan, P.
TITLE Mammalian protein phosphatases
JOURNAL Patent: WO 0146394-A 2 28-JUN-2001;
Sugen, Inc. (US)
FEATURES
source 1. 2732
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.75e-150 Length: 2732
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 6 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x AX180875 (1-2732)
QY 1 MetAlAHieglumeti1egiYThrgInileValThrgluarphleuValAlaleuLeuglu 20
DB 538 ATGGCCCATGAGATGATTGAACTCAATGTTTACTGAGAGGTTGGCTCTGCTGGA 597
QY 21 SerGIYThrgluYsValleuLeuileaspSerArpProPhaValglurYraenThrSer 40
DB 598 AGTGAACCGGAAAGAGTGGCTGTAATGATGACCGGCCATTTGTGAAATCAATATC 657
QY 41 HieileuGlunAlaleanilleanilleanCySerlyLeuMetlySarTargleuIngIn 60
DB 658 CACATTTTGGAAACCAATTAATCACTGCTCAAGCTTATGAGAGCAAGGTTGCAACG 717
QY 61 AspLyValleuileThrgluLeuileglnHieSerAlaYshIlyValAspIleap 80
DB 718 GACAAAGTGTAAATTACAGAGCTCAATCCAGCATTCGCAACATAGAGTTGACATTGAT 777
QY 81 CySerGlnlyeValValValTyraSpGlnSerSerGlnAspValAlaSerLeuSer 100
DB 778 TGCAGTCAGAGTGTGATTTTACGATCAAGCTCCCAAGATGTTGCTCTCTTCA 837
QY 101 AspCyPheleuthrValleuLeuGlylyleuGluYsSerPheAsnSerValHleu 120
DB 838 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGCTTCAACTCTGTTCACTG 897

OY		121	LeuaIagIyglYPhealaglUpheSeserArgCySpheProdlYneUCysgluclyLySer	140
Db		898	CTTCAGAGTGGGTTTGCTGAATTCCTCGTTGTTCCTCGGCCTCTGTAGAAGAAATTC	957
OY		141	ThirdeuValProThrCysIISeSerGlInProCYLeuProValAlAsnIIlegLYProThr	160
Db		958	ACTCTAGTCCCTACCTGSCATTCTCAAGCTTCCTAACCTGTTCCTGTGCCAACAATTGGCCAACC	1017
OY		161	ArgIIeLeuProAsnLeuTYrLeugLYCysglnarAgSPvalleuAsnLysgluleuile	180
Db		1018	GGAATTCCTCCCATCTTAATCTTGCTGCACGACGATGTCCTCAACAAGAGCTGATG	1077
OY		181	GIinglAsnGLyllegLYrValleuAsnAlaserTYrThrCYeProLYeProAsphe	200
Db		1078	CACGAGATGGAGTTGGTTATGTGTTAAATGCCAGCAATACCTGTGCCAAGCTGACATT	1137
OY		201	IleProGLUserHisPheleuArGVALProValaAsnSPserPheCYsglULyslleu	220
Db		1138	ATCCCGAGATCTCAATTCCTGGGTGGCTGTGANTGACAGCTTTTGTGAGAAAATTTTG	1197
OY		221	ProTrPLEuASPLYSerValasPheIIegLUlysAlalyAlaserargLYCYsVAL	240
Db		1198	CGGTGTTGGACAAATCAGTAGAATTTTCATTGAGAAAGCAAAGCTCCAAAGAGATGCTT	1257
OY		241	leuValHisCYleuAlagLYllseSerarGerSalatrnIIeAlaleaIeAtYrIlEmet	260
Db		1258	CTAGTGCACTGTTTGTGCTGGGATCCCGCTCCGCCACCAATCGTATGCTCATCATATG	1317
OY		261	LysArGMeLasPMeSerleuAspGluAlaryArgrgPheValLysglULysArGPcThr	280
Db		1318	AAGAGAGATGACATGCTTTAGATGATMACCTTACAGATTTGTGAAAAAGAAAGCCACT	1377
OY		281	IIseSerProAsnPheAsnPheLUGLYGlnleuLusPYrGluLYSLysIIeLYsAsn	300
Db		1378	AATCTCCAAACTTCATTTCTTGCGCCAACTCCTGACATATGAGAAAGATTAAAGAC	1437
OY		301	GIInthr 302	
Db		1438	CAGACT 1443	
RESULT 7				
LOCUS		AX278461		
DEFINITION		Sequence 1 from Patent WO0177340.		
ACCESSION		AX278461		
VERSION		AX278461.1		
KEYWORDS		. GI:16605915		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE		Duecker, K.		
AUTHORS		1		
TITLE		Identification of a dual specificity phosphatase: dusp-10		
JOURNAL		Patent: WO 0177340-A, 1 18-OCT-2001;		
		MERCK PATENT GmbH (DE)		
FEATURES		Location/Qualifiers		
source		1..3059		
		/organism="Homo sapiens"		
		/mol_type="unassigned DNA"		
		/db_xref="taxon:9606"		
CDS		127..2124		

TP,LP,PC,AD,SA,T,SE,AG,OR,PH,PA,SV,SV,SV,SV,OP,SL,ED,SE,VO,AL,GE,LI,SH,AD,RL,ED
 SN,KR,K,PS,DL,IK,SV,TS,S,AM,MA,LS,EP,FS,SE,DL,EY,Y,K,ST,LL,DN,N,K,CO,FP,VO,ED
 L,SE,OT,PE,T,SP,K,FE,AS,I,PK,I,OL,TA,R,PE,D,SO,GR,KH,SV,TS,SS,GA,OC,SL,PL,SR,SS,ED
 V,ED,N,Y,T,FL,FG,LT,SO,OL,TK,SA,GI,EL,K,M,HS,I,IL,AP,OT,PE,LT,LS,SW,PA,T,SS,HP
 Y,SA,I,Y,G,SA,SV,SV,SV,SV,SO,LP,TC,GD,OV,SV,V,RR,OR,CR,D,AD,SR,SH,ES,PE,FO,K,PK
 R,SC,QM,FE,BS,IM,SE,N,SR,EL,K,VS,G,SS,FS,SM,EL,EY,VS"

Alignment Scores:	
Pred. No.:	5,486-150
Score:	1552.00
Percent Similarity:	99.67%
Best Local Similarity:	99.34%
Query Match:	99.23%
DB:	6
Length:	3059
Matches:	300
Conservative:	1
Mismatches:	1
Indels:	0
Gaps:	0

US-10-029-345A-109_COPY_1_302 (1-302) X AX278461 (1-3059)

QY	1	MetLAHis:sglMetIle:glYThrGlnIleValThrGlnArgLeuValAlaLeuLeuGlu	20
Db	127	ATGGCCCAAGAAATGATATTGGAAACCTCAATATTGTTACTGAGAGGTGGTGGCTGCTGCGAA	181
QY	21	SerGIYThrGluYValLeuLeuLeuIleAspSerArgProPheValGluYrtrsnThrSer	40
Db	187	AGTGGAAACGAAAAAAGTGGCTGCTAAATTGATGACCGGCCATTGTGGAAATACATACATCC	244
QY	41	HisIleLeuGluAlaIleAsnIleAsnCysSerValLeuMetLeuYArgYrghLeuGlnGln	60
Db	247	CACATTTTGGAAAGCCATTAAATCACTGCTCCAGCTTATATGACGAAAGGTGGCAACG	306
QY	61	AspIysValIleuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp	80
Db	307	GACAAAGCTTTAATTACAGAGCTCATCTCCAGCATTCAGCCAAAACAATAGATTGACATTGAT	366
QY	81	CysSerGlnYsValValIlyrAspGlnSerSerGlnAspValAlaSerLeuSerSer	100
Db	367	TGCAGTCAAGAGGTGTAGTTTACGATCAAAAGTCCCAAGATGTTGGCTCTCTCTCTTCA	426
QY	101	AspCysPheLeuThrValIleLeuLeuGlyLysLeuGlnYsSerPheAsnSerValHisLeu	120
Db	427	GACGTGTTTTCACCTGATCTTGGGTAAACGTGGAAGAGACCTTCACTGTTCACTG	486
QY	121	LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlnGlyLysSer	140
Db	487	CTTCGAGTGGGTTTGCTAGATTCCTCTCGTGTGTTCCCTGGCCCTGTGGAAGAAATCC	546
QY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	547	ACTCTAGTCCCTACCTGCATTTCTCAGGCTTGCTTACTGTTGCCAATTTGGGCCAAC	606
QY	161	ArgIleLeuProAsnLeuYrIleuGlyCysGlnArgAspValLeuAsnYsGluLeuIle	180
Db	607	CGAATTTTCCCAATCTTTATCTTTGGCCGCAAGAGATGCTTCAACAAGAGAGCTGAG	666
QY	181	GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerYrIthrCysProLysProAspPhe	200
Db	667	CAGCAGAGATGGGATTGGTTATGTGTTAAATCCGACAAATCCTGTCCAAAGCTGACTT	726
QY	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnYsIleLeu	220
Db	727	ATCCCCGAGTCCATTTCTCGTGCTGTGATGACAGCTTTTGTGAGAAATTTTG	786
QY	221	ProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal	240
Db	787	CGGTGGTTGGACAAATACGTAGATTTTATGAGAAAGCAAAAGCCCAATGATGATGTGT	846
QY	241	IleuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaYrIleMet	260
Db	847	CTAGTGCACGTTTATAGCTGGGATCTCCCGCTCCGCCACATGCGCTATGCGCTTACATCA	906
QY	261	LysArgMetAspMetSerLeuAspGluAlaYrArgPheValLysGluLysArgProThr	280
Db	907	AAGAGAGTGGACATGCTCTTATAGTGAAGCTTACAGATTTGTGAAGAAAAAAGATTTCT	966

Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.F., Schat, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Butler, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marisina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Brownstein, M.U., Ussidi, T.B., Toshiyuki, S., Carmichael, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

TITLE	JOURNAL	MEDLINE PUBMED REFERENCE	AUTHORS	REMARK COMMENT
Abraham, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilly, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonfard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywiec, M.I., Skalka, U., Smilans, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Katta, M.R.	Submitted (23-DEC-2002)	National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web Site: http://www.sngc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Series: IRAK Plate: 88 Row: a Column: 17 This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction.				
location/Qualifiers				
1. .3284				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="MGC:50665 IMAGE:4400399"				
/tissue_type="Duodenum, adenocarcinoma"				
/clone_lib="NIH MGC_88"				
/lab_host="DH10B"				
/note="vector: pCMV-SPORT6"				
1. .3284				
/gene="DUSP16"				
/note="synonyms: KIAA1700, MKP-7, MKP7"				
/db_xref="LocustID:80824"				
/db_xref="MIM:6071175"				
332_ .2320				
/codon_start=1				
/product="DUSP16 protein"				
/protein_id="AA42101.1"				
/db_xref="GI:27469789"				
/db_xref="LocustID:80824"				
/translation="MAHEMTGQIVTERLVALLSEGTKEVLLIDSPFVYNSHILE AINNSKLMKRRQODPKVLTIELIGSAKHKDIDSGQVYVTDSSQVASTSDSC FLTVLGLKFKSFNSVHLLAGFARFEPSPGGLCEGSTLVPTISQPCPLVNIIGPT RLIPNLVIGQDRVNLKELMQNGIIGFVLANSTCPKPDFIPESHFLRPVNDSCFKK ILPLDSDVDFIEKAKNSGCVVHCAIGASRSATLILAIYIKPMDSLDEAYFVE KRPGLISNENFGLDIDYEKKIKNOTGASGPKSLKLHLKEPMRVPVAYSEGQKE TLPSPRDSATSEAGORVHPASVPSVOPSLIENSLVLAOGLSHLSADRLIEDSNK LKSFSIDIKSVSYASMASJIGFSSSEDALEYKSTLTLDGNTLCQSPVOELSE QTEPTSDKEBASIPKRLQTFARPDSOSKRLHSVRISSGFAQNSLSPILHRSQVED NHTSPFLGISTSQOHLTKSAGLKGWHDILAPQRTSLTISWYFAETSHFYS SATYGGSAVSAYSQGLPTCGQVYVSRVRRORSRADSADRSWHESPEKQFKRS				

misc_feature	380..730	COMERGEISMENSRREELKAVQSGSPSGSMELLIEVS"
	/note="KHOD; Region: Rhodanese Homology Domain"	
misc_feature	803..1222	/db_xref="CDD:smat00450"
	/note="DSPC; Region: Dual specificity phosphatase, catalytic domain. Ser/Thr and Tyr protein phosphatases. The enzyme's tertiary fold is highly similar to that of tyrosine-specific phosphatases, except for a 'recognition region"	
	/db_xref="CDD:pfam00782"	
ORIGIN		
Alignment Scores:		
Pred. No.:	6e-150	Length: 3284
Score:	1552.00	Matches: 300
Percent Similarity:	99.67%	Conservative: 1
Best Local Similarity:	99.34%	Mismatches: 1
Query Match:	99.23%	Indels: 0
DB:	9	Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x BC042101 (1-3284)		
QY	1	MetAlAHISGLMETILLEGlyThrGlnIleValIThrGluArgLeuValAlaLeuLeuGlu 20
DB	332	ATGGCCCAATGAGATGATATTGGAACCTCAAAATTTGTACTGAGAGGTGGTGGCTCTCGTGGA 391
QY	21	SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB	332	AGTGAACGGAAGAAAGTGGCTCTGTAATTATTAAGCCGCGCATTTGTGGAATACATATCATCC 451
QY	41	HisIleLeuGluAlaIleAsnIleAsnCySerIlySleuMetIlyArgArgLeuGln 60
DB	452	CACATTTTGGAAAGCATTAAATATCAACTGCTCCAAAGCTTAATGAACGGAAGTTGCAACAG 511
QY	61	AspIlyValIleuIleThrGluLeuIleGlnHisSerAlaIlyHisIlyValAspIleAsp 80
DB	512	GACAAAGGTTAATTAACAGAGCTCATTCAGACATTCAGCAAAACATTAAGGTTGACATTGAT 571
QY	81	CysSerGlnIlyValIleValIlyTrpAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB	572	TGCAGTCAGAAAGTTGTGATTTAGATCAAAAGCTCCAAAGATGTTGCTCTCTCTTCA 631
QY	101	AspCysPheLeuThrValLeuLeuGlnIlyIlySleuGlnIlySerPheAsnSerValHisLeu 120
DB	632	GACTGTTTTCACACGATCACTCTGGGTAACTGGAGAAAGACCTCACTCTTCACTG 691
QY	121	LeuAlaGlyIlyPheAlaGluPheSerArgCysPheProGlyLeuCyGlnIlyIlySer 140
DB	632	CTTCGAGGTGGGTTTGTGAGGTTCTCTGTTGTTTCCCTGGGCTCTGGAAGGAAATCC 751
QY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB	752	ACTGAGTCCTTACCTGATTCCTCAGGCTTGCTTACTGTTGCGCAATTTGGAGCAAC 811
QY	161	ArgIleLeuProAsnLeuIlyLeuGlnIlyCysGlnArgAspValIleAsnIlyGlnLeuIle 180
DB	812	GGAATCTTCCCAATCTTATCTTGCGCTGGCCAGGAGATGTCCTCAAAAGGAGCTGATG 871
QY	181	GlnGlnAsnGlnIlyIleGlyTrpValLeuAsnAlaSerTrpThrCysProIlyProAspPhe 200
DB	872	CAGCAGATGAGGATTTGGTTATGTGTAAATCCAGCAATATACCTGCCAAAGCCTGACCTT 931
QY	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIlyIleu 220
DB	932	ATCCCGCAGTCTCATTTCTCCGCTGGCTCTGATGACACTTTTGGACAAATAATTTTG 991
QY	221	ProTrpLeuAspIlySerSerValAspPheIleGlnIlyAlaIlyAlaSerAsnGlnIlyVal 240
DB	992	CCGAGGTTGGCAATCAATGATGATTTCAITTAGAAAGCAAAAGCTCCAAATGAGATGTGT 1051
QY	241	LeuValHisCysLeuAlaGlnIlyIleSerArgSerAlaThrIleAlaIleAlaIlyTrpLeu 260

Db 1052 CTAGTGCAGCTGTTAGCTGGAGATCTCCCGCTCCGCCACCATGCTATGCGCTACATCATG 1111
Qy 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLysArgProThr 280
Db 1112 AAGAGATGAGACATGCTTTAGATGAGAGCTTACAGATTGTGTGAAGAAAAAGACCTACT 1171
Qy 281 ILeSerProAspPheAspPheLeuGlyIleLeuLeuAspTyrGluLysIleLysAsn 300
Db 1172 ATATCTCCAAACTTCATTTTCTGGGCCAACCTCTGACTATGAGAAAGATTAAAGAAC 1231
Qy 301 GlnThr 302
Db 1232 CAGACT 1237
RESULT 10
AX441210 3496 bp DNA linear PAT 28-JUN-2002
LOCUS Sequence 1 from Patent WO226997.
DEFINITION AX441210
ACCESSION AX441210
VERSION AX441210.1 GI:21665766
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Luche, R.M. and Wei, B.
TITLE Dep-16 dual-specificity phosphatase
JOURNAL Patent: WO 0226997-A 1 04-APR-2002;
Cepcyr, Inc. (US)
FEATURES
source location/Qualifiers
1..3496 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6..5e-150 Length: 3496
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 6 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x AX441210 (1-3496)
Qy 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
Db 562 ATGGCCCATGAGATGATGAGACTCANAATTGTACTGAGAGGTGGCTCTGCTGGAA 621
Qy 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db 622 AGTGGAAAGCGAAAGAGCGCTGCTAATTGATGAGCGGCAATTGTGGAAATACATACATCC 681
Qy 41 HisIleLeuGluAlaIleAsnIleAsnCysSerLysLeuMetLysArgIleGlnGln 60
Db 682 CACATTTTGGAGCCATTATATATCACTGCTCCAGCTTATGAAAGCGAGTTGCAACAG 741
Qy 61 AspLeuValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 742 GACAAAGTGTAAATTAACAGCTCATCCAGCATTCAGCGAAACATTAAGGTGACATTGAT 801
Qy 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 802 TGCAGTCAGAGAGTTGATGTTACGATCAAAAGCTCCCAAGATTTGCTCTCTTCA 861
Qy 101 AspCysPheLeuThrValLeuLeuGlyLysGluLysSerPheAsnSerValHisLeu 120
Db 862 GACTGTTTCTCACTGACTCTGGGGTAACTGGAAGAGCTTCAACTCTGTTCACTCG 921
Qy 121 LeuAlaGlyCysPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140

Db 922 CTTCGAGGAGGTTTGCTGATGTTCTCTGTTTTCCTTGAGCTCTGTGAGGAAAAATCC 981
Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 982 ACTCTAGTCCCTTACCTGCTGATTTCTCAGCTTCTTACCTGTTCCAACTTTGGCCAAAC 1041
Qy 161 ArgIleLeuProAspLeuLysIleGlyCysGlnArgAspValLeuAsnLysGluLeu 180
Db 1042 CCAATTTCTCCAAATCTTTATCTGGCTCCAGCGAGATGCTCTCAACAGAGCTGATG 1101
Qy 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
Db 1102 CAGCAAGATGGATTTGTTATGTTAAATGACAGAAATACCTGTCAAAGCTGACTTT 1161
Qy 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
Db 1162 ATCCCGGATCTTCATTTCTGCTGGCTGCTGAAATGACAGCTTTGTGAGAAATTTTG 1221
Qy 221 ProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
Db 1222 CCGTGGTGGACAAATCACTAATTCATTGAGAAAGCNAAGCTTCCAAATGATGTGT 1281
Qy 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db 1282 CTAGTGCAGCTGTTTAGTGGATCTCCGCTCCGCCACCATGCTATGCGCTACATCATG 1341
Qy 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLysArgProThr 280
Db 1342 AAGAGATGAGACATGCTTTAGATGAGAGCTTACAGATTGTGAAAGAAAAAGACCTACT 1401
Qy 281 ILeSerProAspPheAspPheLeuGlyIleLeuLeuAspTyrGluLysIleLysAsn 300
Db 1402 ATATCTCCAAACTTCATTTTCTGGGCCAACCTCTGAGACTATGAGAAAGATTAAAGAAC 1461
Qy 301 GlnThr 302
Db 1462 CAGACT 1467
RESULT 11
AB052156 3521 bp mRNA linear PRI 18-OCT-2001
LOCUS Homo sapiens MKP-7 mRNA for MAPK phosphatase-7, complete cds.
DEFINITION AB052156
ACCESSION AB052156
VERSION AB052156.1 GI:13548676
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Masuda, K., Shima, H., Watanabe, M. and Kikuchi, K.
TITLE MKP-7, a novel mitogen-activated protein kinase phosphatase,
functions as a shuttle protein
JOURNAL J. Biol. Chem. 276 (42), 39002-39011 (2001)
MEDLINE 21486429
PUBMED 11489891
REFERENCE 2 (bases 1 to 3521)
AUTHORS Masuda, K., Shima, H. and Kikuchi, K.
TITLE Direct Submision
JOURNAL Submitted (05-DEC-2000) Kouhei Masuda, Institute for Genetic
Medicine, Division of Biochemical Oncology and Immunology, Kita-ku
Kita15-5-1, Nishi-ku, Sapporo, Hokkaido 060-0815, Japan
(E-mail: kou@imm.hokudai.ac.jp, Tel: 81-11-706-5536,
Fax: 81-11-707-6839)
FEATURES
source location/Qualifiers
1..3521 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/cell_line="Jurkat"
1..3521 /gene="MKP-7"

CDS

564.. 2561
/gene="MKP-7"
/codon_start=1
/product="MAPK phosphatase-7"
/protein_id="BAB40814.1"
/db_xref="GI:13548677"
/translation="MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILE
AININCSKLMKRLQODKVLITELIHSAAHKVVIDSOXVVYDSSODVASLSDC
FLTVLLGKLEKSFNSVHLIAGFAEPRCPGLCEGKSTLVPCTISQPCLPVNIIGPT
FLTVLLGKLEKSFNSVHLIAGFAEPRCPGLCEGKSTLVPCTISQPCLPVNIIGPT
FLTVLLGKLEKSFNSVHLIAGFAEPRCPGLCEGKSTLVPCTISQPCLPVNIIGPT
ILPWLKSVDFIEKAKASNGCVLHCLAGLSRSATIAIAYIMKMDSLDEAYRVE
KRPTISPNFNLGQLIDYEKKIKNOCTAGSGKSLKILHIEKPEPVPAYSEGGQSE
TPLEPCADATSEAGQRPVHPASVSPVSPSLDESPLYOALSGLHISADRLSD
SNKIKSFSLIDIKSVSASMAASLHGFSSSEDALEYKSPSTLIDGNKLCQSPVDS
LSBOTPTSPDKERASTPKLQOTARPDSOSKRLHSVTSSTSGTASRLSLPLHSGS
VEDNYHTSFLPGISTSQOHLTKSAGLHKMGHSIDILAPQSTPSLTSSWYFATESSHF
YSASATYGGASAYSACSQPLTCGDQVYSVRRROKPSDRADSRHSWHESPEKQPK
RRSCOMEFGESIMSENRSRBELGKVGSSQSSFSGSMETIEVS"

ORIGIN

Alignment Scores:
Pred. No.: 6.55e-150 Length: 3521
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AB052156 (1-3521)

QY 1 MetAlahisglumellleglythrglnilevalthrgluargleuvalaleuleu 20
DB 564 ATGGCCCATGAGATGATTGGAACCAATTGTTACTGAGAGGTTGGTGGCTCTGGGAA 623
QY 21 SerGlythrgluysvalleuleuileaspserrpophelvalglutyrasnthser 40
DB 624 AGTGAACGGAAAAAGTCCTGCTATTGATAGCCGGCCATTGGGAATTAACAATACC 683
QY 41 Hisileuagluajalleasnilleasnyserrlysluemetlysarargleuagln 60
DB 684 CACATTGGAAAGCCATTAAATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACG 743
QY 61 AsplysValleuilethrgluuileglnhiserralysrhislyvalasp 80
DB 744 GACAAAGGTTAATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 803
QY 81 CysSerGlnlysvaivaivaityrAspGlnserSerGlnasppvalalaserleuSer 100
DB 804 TGCAGTCAGAAAGGTTGATGTTACATCAAAAGCTCCAAAGATGTTGCTCTCTCTTCA 863
QY 101 AspCysPheleuThrValleuleuGlylysleuGlyysSerPheAsnSerValHisleu 120
DB 864 GACTGTTTTCACCTGACTCTCTGGGTAACTCGAAGAAAGACTTCACTCTGTTCACTG 923
QY 121 LeuAlaglygylphealagluhpeserargCysPheProGlyLeuGlyGlylysSer 140
DB 924 CTTCAGAGGTGGTTGCTGAGATTCCTCGTTGTTCTCCCTGGCCTCTGAAAGAAATCC 983
QY 141 ThrleuValProThnCysIleSerGlnProCysleuProvalAlaAsnIleGlyProthr 160
DB 984 ACTCTAGTCCTACCTGCAATTCCTGAGCTTGCTTAACCTGTTGCAAAATGGGCAACC 1043
QY 161 ArgIleleuProAsnleuThrleuGlyCysGlnargAspValleuAsnlysglyleu 180
DB 1044 CGAATTCCTCCAACTTATCTTGGCTGCCAGCAAGATGCTCTCAACAAGAGCTGATG 1103
QY 181 GlnGlnaenglylleglytyrValleuAsnAlaserTythrCysProlyPheAspPhe 200
DB 1104 CAGCAGATGGGATGTTAGTGTAAATGCAACAATACCTGCTCCAAAGCTGACCTT 1163
QY 201 IleProGlnUserHisPheleuargValaProvalaAsnAspSerPheCysGlnlyslleu 220
DB 1164 ATCCCGAGTCTCATTTCTCTGGTGGTGGCTGTAATGAACGCTTTTGTAGAAATTTTG 1223

QY 221 ProTripleuAspySerValaaspheilegulyvalalyalaseranglyCysVal 240
DB 1224 CCGTGTGGACAAATCAGTAGATTTCATTGAGAAAAGCAAAAGCTCCATGATGATT 1283
QY 241 LeuValHisCysleuAlaglylleserargserAlaThrIleAlaIleAlaTyrlleMet 260
DB 1284 CTAGTGACACTGTTAGCTGGATCTCCCGCTCCGCCACCAATCCGTATCGCTTATCATCATG 1343
QY 261 LysArgMetAspSerleuAspGluAlaTyArgPheValIlysglylysArgProThr 280
DB 1344 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTGTGAAAAGAAAAAGACTTACT 1403
QY 281 IleSerProAsnPheAsnPheleuGlyGlnleuAspTyrglylyslyllyAsn 300
DB 1404 ATATCTCAAACTTCATTTTCTGGGCCCAACTCCTGAGATATGAGAAAGATTAAGAAC 1463
QY 301 GlnThr 302
DB 1464 CAGACT 1469

RESULT 12
AX260340
LOCUS AX260340 3544 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 1 from Patent WO0173059.
ACCESSION AX260340
VERSION AX260340.1 GI:16509303
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Meyers, R.A.
TITLE 1
JOURNAL 36692 and 21117: dual specificity phosphatase molecules and uses
FEATURES
source
LOCATION/Qualifiers
1..3544
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
589..2586
/note="unamed protein product"
/codon_start=1
/protein_id="CAD10218.1"
/db_xref="GI:16509304"
/translation="MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILE
AININCSKLMKRLQODKVLITELIHSAAHKVVIDSOXVVYDSSODVASLSDC
FLTVLLGKLEKSFNSVHLIAGFAEPRCPGLCEGKSTLVPCTISQPCLPVNIIGPT
FLTVLLGKLEKSFNSVHLIAGFAEPRCPGLCEGKSTLVPCTISQPCLPVNIIGPT
ILPWLKSVDFIEKAKASNGCVLHCLAGLSRSATIAIAYIMKMDSLDEAYRVE
KRPTISPNFNLGQLIDYEKKIKNOCTAGSGKSLKILHIEKPEPVPAYSEGGQSE
TPLEPCADATSEAGQRPVHPASVSPVSPSLDESPLYOALSGLHISADRLSD
SNKIKSFSLIDIKSVSASMAASLHGFSSSEDALEYKSPSTLIDGNKLCQSPVDS
LSBOTPTSPDKERASTPKLQOTARPDSOSKRLHSVTSSTSGTASRLSLPLHSGS
VEDNYHTSFLPGISTSQOHLTKSAGLHKMGHSIDILAPQSTPSLTSSWYFATESSHF
YSASATYGGASAYSACSQPLTCGDQVYSVRRROKPSDRADSRHSWHESPEKQPK
RRSCOMEFGESIMSENRSRBELGKVGSSQSSFSGSMETIEVS"

CDS

ORIGIN

Alignment Scores:
Pred. No.: 6.61e-150 Length: 3544
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AX260340 (1-3544)

QY	1	20
QY	1	MetalAaHsGluMetIIGIYThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu
Db	589	ATGGCCCATGAGATCATTTGGAACTCAAAATTTGTTACTGAGAGGTTGTGGCTCTGCTGANA
QY	21	SerGIYThrGluValLeuLeuLeuIleAspSerArgProPheValGluTyraThrSer
Db	649	AGTGGAAACGAAAAAGTCTGCTGTAATTATATAGCCGGCATTTGTGGAATPACAATACATCC
QY	41	HisIleLeuGluAlaIleAsnIleAsnCyserIleuMetIleValArgLeuGlnGln
Db	709	CACATTTTGGAAAGCATTAAATATCAACTGCTCCAACTTAATGAACGGAAGGTTGCACAG
QY	61	AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleHisIleValAspIleAsp
Db	769	GACAAAGTGTTAATTACAGACTCATCAAGCACTTAGGAAACATTAAGTTGACATTTAT
QY	81	CysSerGlnIleValIleValTyraAspGlnSerSerGlnAspValAlaSerLeuSerSer
Db	829	TGCAGTCAAGAAAGTTGTATGTTACATCAAGCTCCAAAGATGTTGCTCTCTCTTCA
QY	101	AspCysPheLeuThrValLeuLeuGluIleIleLeuGluIleIleSerPheAsnSerValHisLeu
Db	889	GACTGTTTTCACACTGACTCTTGGGGTAAACGTGGAAGAGACTTCAACTCTGTTCACTG
QY	121	LeuIleGluIlePheAlaGluPheSerArgCysPheProGluIleCysGluGluIleIleSer
Db	949	CTTGAGAGTGGTGTGCTGAGTCTCTGTGTGTTCCCTGGGCTCTGTGAAGGAAATCC
QY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGluProThr
Db	1009	ACTCTAGTCCCTACCTGCATTTCTCAGCCTTGCTTACCTGTGGCCAACTTGGGCCAAC
QY	161	ArgIleLeuProAsnLeuTyreLeuGluIleCysGlnArgAspValLeuAsnIleGluLeuIle
Db	1069	CGAATTTCTCCCAATCTTATCTTGCTGGCTGCCAGCAGATGTCCTCAACAGAGCTGATG
QY	181	GlnGlnAsnGluIleGluIleTyraIleValLeuAsnAlaSerIleTyraCysProIleProAspPhe
Db	1129	CAGCAGAAATGGATGGATTATGTGTAAATGCCACCAATACCTGTCCAAAGCTGACATTT
QY	201	IleProGlnSerHisPheLeuArgValProValaAsnAspSerPheCysGluIleValLeu
Db	1189	ATCCCCGAGTCTCATTTCTCTGGGTGTGCTGTGAAATGACAGCTTTTGTGAGAAATTTTG
QY	221	ProTrpLeuAspIleSerValAspPheIleGluIleValAlaIleAsnIleCysVal
Db	1249	CCGTGGTTGGACAATCAATGATTTTCAATGAGAAAGCAAAAGCTCCAAATGAGATGTTT
QY	241	LeuValHisCysLeuAlaGluIleIleSerArgSerAlaThrIleAlaIleAlaIleIleIle
Db	1309	CTAGTGCAGCTGTTTGTGCTGGGATCTCCGGCTCCGCCACATCGCTATGCGCTACATCATG
QY	261	LysArgMetAspMetSerLeuAspGluAlaTyraArgPheValIleGluIleArgProThr
Db	1369	AAGAGGATGGACATGCTTTTATGATTAACCTTACAAATTTGGAAGAAAAAGACTCATCT
QY	281	IleSerProAsnPheAsnPheLeuGluIleGlnLeuAspTyraGluIleIleIleIleIleIle
Db	1429	ATATCTCCAAACTTCAATTTCTGGGCCAACTCTCGAGCATATGAGAAAGAAATTAAGAAC
QY	301	GlnThr 302
Db	1489	CAGACT 1494
RESULT 13		
AF506796	AF506796	3566 bp mRNA linear PRI 30-OCT-2003
LOCUS	DEFINITION	Homo sapiens dual specificity phosphatase 16 (DUSP16) mRNA, complete cds.
ACCESSION	AF506796	
VERSION	AF506796.1	GI:25573087
KEYWORDS	Homo sapiens (human)	
SOURCE		

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 3566)
TITLE	Hoornaert, I., Marynen, P., Goris, J., Sciote, R. and Baens, M.
JOURNAL	MAPK phosphatase DUSP16/MKP-7, a candidate tumor suppressor for chromosome region 12p12-13, reduces BCR-ABL-induced transformation
REFERENCE	Oncogene 22 (49), 7728-7736 (2003)
AUTHORS	14586359
JOURNAL	2 (bases 1 to 3566)
REFERENCE	Hoornaert, I., Marynen, P. and Baens, M.
AUTHORS	Submitted 126-APR-2002) Department for Human Genetics-Flanders
TITLE	Direct Submission
JOURNAL	Interuniversity Institute for Biotechnology (VIB), Katholieke
REFERENCE	Universiteit Leuven, Herestraat 49, Leuven B-3000, Belgium
FEATURES	Location/Qualifiers
source	1..3566
gene	/organism="Homo sapiens"
CDs	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosome="12"
	/map="12p13-p12"
	/tissue_type="bone marrow"
	1..3566
	/gene="DUSP16"
	/note="synonym: MKP7"
	633..2630
	/gene="DUSP16"
	/EC_number="3.1.3.48"
	/EC_number="3.1.3.16"
	/codon_start=1
	/product="dual specificity phosphatase 16"
	/protein_id="AA075120.1"
	/db_xref="GI:25573088"
	/translation="MAHEMIGTQIVTERLVALLHSGTEKYLIDSRPFVRYNTSHILE
	AINUNISAKMKRRLQODKVLITELIDHSAHKYDIDCSQVVVYDQSDQVSLASDPC
	FLTVLVLGKLEKSRNSHLLAGFAERSRCFGLCEGKSTIVPTICISQPLPVANIGPT
	RIILPNLYLGGQDRLNKEMLQNGIGVILNASTCTCPEDTIPESHFLRVVNSFCER
	ILWPLDKSVDFIEKARASNGCVLHCLAGISRSATLAIATIMKRMVMSLDEARFVKE
	KRRTLSINFNFLQGLDYEKKIKINQGTASGSPKSKLILHEKENEPVPAVEGQKSE
	TPSLPCADATSEAGAGQRPVHPASVPSVQPSLLESDPLVQALSGHLSDRLIED
	SNLKTSFSLDITKSVSASMASLHSGSSSDALRYKSTSLDGNKTLICQSPQVSE
	LSQCTPETSPPDKKEBASIPKRLQTPARPSDSQSKLHVSRTSSGTQKSLSLPHRSS
	VENNYHTSFLFGLSTSQQLHTKSAGLGKGMHSDILAPQSTSLTSSWTFATESSHF
	YSSALVIGSGASYSAYSCSQLPTCGDQVSVVRROKRPSPADSRSMHESRPEKQPK
	RRSQWFEGESIMSENRSRELKGVGSSSGSMELIEVS"
ORIGIN	
Alignment Scores:	6.66e-150 Length: 3566
Pred. No.:	1552.00 Matches: 300
Score:	99.67% Conservative: 1
Percent Similarity:	99.34% Mismatches: 1
Best Local Similarity:	99.23% Indels: 0
Query Match:	Gaps: 0
DB:	9
US-10-029-345A-109_COPY_1_302 (1-302) x AF506796 (1-3566)	
Qy	1 MetcLahtagLwcttLleglYthnglnlLevalThthgluargLeuValalaleuLeuGlu 20
Db	633 ATGGCCCATAGAGATGATGGAACTCAAAATGTTACTGAGAGGTGGCGCTCTGCTGGAA 692
Qy	21 SerclYthrgluYvalLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db	693 AGTGGAAACGAAAAAGCTGCTGCTATGATGACCGGCGCATTTGTGGAAATCAATCACTCC 752
Qy	41 HislleLengluValaleaaniLeaenCySeSerlySeuWetlybArqLeuGln 60
Db	753 CACATTTGGAGGCATTATATCAACTGCTCCAAAGCTTATGAAAGCAAGTGGCAACAG 812
Qy	61 AsplybValleuLeuLeuThrgluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Db	813 GACAAATGTTAATTAACAGAGCTCATTCAGCAATTCAGCGAAACATTAAGATTGACATTGAT 872

QY	81	CysSerGlnIlysValValValIlyrAspGlnSerSerGlnAspValAlaSerLeuSerSer	100
Db	873	TGcAGTCAGAAAGGTGTAGTTTACGATCAAAAGCTCCAAATGTTGCTTCTCTCTTCA	932
QY	101	AspCyAspHeuThrValIleuLeuGlyLysLeuGlnIlySerPheAsnSerValIhIleu	120
Db	933	GACTGTTTTCTCACTGACTCTTGCGGTAACTGGAGAAAGAGCTTCAACTCTGTTACCTG	992
QY	121	LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGlnGlyLysSer	140
Db	993	CTTGCAAGTGGGTTTGGCTGAGTTCTCTCGTTGTTTCCCTGGCCCTGTGTGAAGAAAAATCC	1052
QY	141	ThrIleuValIProthrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	1053	ACTCTAGTCTCCACTCTCATTTCTCAGCTCTTGCTTACTCTGTGGCAACATTGGGCCAAC	1112
QY	161	ArgIleLeuPProAsnLeuIlyrLeuGlyCysGlnArgAspValIleuAsnIlysgluIleuIle	180
Db	1113	CGAATTTCTTCCATCTTTATCTTGGCTGCAGAGAGATGCTTCAACAAGAGAGCTGATG	1172
QY	181	GlnGlnAsnGlyIleGlyIlyrValIleuAsnAlaSerIlyrThrCysProLysProAspPhe	200
Db	1173	CACGACAGATGGATTTGGTTATGTGTAAATGCCAGCAATACCTGTGCCAAAGCCCTGACTTT	1232
QY	201	IleProGlnSerHisPheLeuArgValIProValAsnAspSerPheCysGlnIlyIleu	220
Db	1233	ATCCCGAGTCTCATTTCTTCGCTGCTGCTGTGATGACAGCTTTGTGGAAAAATTTTG	1292
QY	221	ProIlyrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal	240
Db	1293	CCGGTGTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCCTCAATGAGATGTGTT	1352
QY	241	LeuValHisCysLeuAlaGlyIleSerArgSerIleMetIleAlaIleIyIleMet	260
Db	1353	CTAGTGCACTGTTTAACTGGAGATCTCCGCTCCGACACATGCGCATGCTTCAATCATG	1412
QY	261	LysArgMetAspMetSerLeuAspGlnAlaIlyrArgPheValIlysgluLysArgProThr	280
Db	1413	AAGAGATGACACACTCTTTAGATCAACTTACAGATTGTGAAAAAGAAAAAGACTTACT	1472
QY	281	IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIlyrGluLysIlyIleLysAsn	300
Db	1473	ATATCTCAAACTTCAATTTTCTGGGCCAACTCTCGACTATGAGAAAGATTAAAGAAC	1532
QY	301	GlnThr 302	
Db	1533	CAGACT 1538	
RESULT 14			
AX374994	AX374994	3766 bp	DNA
LOCUS	AX374994	3766 bp	DNA
DEFINITION	Sequence 17 from Patent WO0210363.		linear
ACCESSION	AX374994		
VERSION	AX374994.1	GI:19169826	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE			
AUTHORS	1		
	Wang,Y.T., Elliott,V.S., Ramkumar,J., Yao,M.G., Burford,N.,		
	Tang,Y.E., Stewart,E.A., Gandhi,A.R., Patterson,C., Lee,B.A.,		
	Hafellia,A.J., Lu,D.A., Tribouley,C.M., Griffin,J.A., Baughn,M.R.,		
	Yue,H., Warren,B.A., Nguyen,D.B. and Wallia,N.K.		
TITLE	Protein phosphatases		
JOURNAL	Patent: WO 0210363-A 17 07-FEB-2002;		
	Incyte Genomics, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..3766		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		

ORIGIN	/note="Inocyte ID No: 7480570CB1"
Alignment Scores:	7.14e-150
Pred. No.:	1552.00
Score:	99.67%
Percent Similarity:	99.34%
Best Local Similarity:	99.23%
Query Match:	6
DB:	Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x AX374994 (1-3766)	
QY	1 Met1ahsgumet11eglythrgn111eValThrgluargLeuVal1aleuLeuglu 20
DB	538 ATGGCCATGAGATGATTGGAATCCTCAATTTGTTACTGAGAGTGGTGCTGCTGGAA 597
QY	21 SerglythrgluysValleuLeu11easpSerargPropheValGluYrAnthrSer 40
DB	598 AGTGAACGGAAGAAAGTGTGCTCAATGATAGCCGCGCATTTGGMATCAATACATCC 657
QY	41 His11eLeuglu1a11eLeu111easnCyser11yLeuMet1ysArg1eug1ng1n 60
DB	658 CACATTTGGAGGACATTAATCACTGCTCCAGCTTAAAGAGCAAGCTTGCACAG 717
QY	61 AspyVal11eui1etThrgluLeu11eg1nHisSer11a1yshi1yVal1asp11eap 80
DB	718 GACAAAGTGTAA1TTACAGAGCTATCCAGATTTACAGCAAAACATAAGTTGACATTGAT 777
QY	81 CysSer11yVal1Val1Val1Yr1asp11nSerSer11naspVal11aSer11eSerSer 100
DB	778 TGCAGTCAGAGGTTGATGATTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 837
QY	101 AspySpheLeuThrVal11eLeu11y11y11ySer11ySphe11nSerVal11His11eLeu 120
DB	838 GACTGTTTTCACACTGATCTTCTGG11AACTGAGAAAGACCTTCAACTCTGTTCACTG 897
QY	121 Leu11a11y11yPhe11a11y11yPhe11ySer11yCysp11hePro11yLeu11y11ySer 140
DB	898 CTTCAGAGTGGGTTTGGTGAATCTTCTGTTCTTCCCTGGCCCTGTGAAGGAAATTC 957
QY	141 Thr11eVal11ProThr11y11Ser11nPro11y11eLeu11ProVal11a11asn11eg11yProThr 160
DB	958 ACTGATGCTCCACTGCTGATTTCTCAAGCTTGTCTTACCTGTTGCAACATTTGGGCAAC 1017
QY	161 Arg111eLeu11Pro11nLeu11y11y11yCys11n11yArg11yVal11e11e11y11y11eLeu 180
DB	1018 CGAATTCCTCCCAATCTTATCTTGGTGGCTGCAAGCATGCTCTCAACAGAGACTATG 1077
QY	181 G11n11a11nG11y11eg11y11yVal11e11y11a11a11Ser11yThr11y11Pro11y11Pro11ySphe 200
DB	1078 CAGCAGAA11GGGAT11GGTATGTTAATCCAGCAATACCTGTCCAAAGCCGACATTT 1137
QY	201 11ePro11y11Ser11His11Phe11y11yVal11ProVal11a11n11y11Ser11yPhe11y11y11eLeu 220
DB	1138 ATCCCGGAGTCTCATTTCTCTGGTGTGCTGTGATGACACCTTTTGGAGAAAATTTTG 1197
QY	221 Pro11Th11e11y11y11ySer11yVal11y11y11y11y11y11y11y11y11y11y11y11y11y 240
DB	1198 CCGTGGTGGCAAAATTCAGATGATTTCTTGAAGAAAGCAAGCTCCAAATGAGATGCTT 1257
QY	241 LeuVal11His11y11e11y11y11e11y11y11e11y11y11e11y11y11e11y11e11y11e11y 260
DB	1258 CTAGTGCATGTTTAGTCTGGAGTCTCCGCGCTCCGCAATCTGCTATGCTTACATCTATG 1317
QY	261 11y 280
DB	1318 AAGAGGATGAGATGCTTTTACATGATAGCTTACATATTTTGGAGAAAGAAAAGACCTTCT 1377
QY	281 11e11y 300
DB	1378 ATATCTCCAAACTTCAATTTTCTGGGCAACCTCTGGACTATGAGAAAGAAATTAAGAAC 1437

QY 301 GlnThr 302
DB 1438 CAGACT 1443

RESULT 15
BD171157
LOCUS BD171157 4790 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel gene and protein encoded thereby.
ACCESSION BD171157
VERSION BD171157.1 GI:27876969
KEYWORDS WO 02052005-A/13.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 4790)
AUTHORS Ohara,O., Nagase,T. and Nakajima,D.
TITLE Novel gene and protein encoded thereby
JOURNAL Patent: WO 02052005-A 13 04-JUL-2002;
KAZUSA DNA RESEARCH INSTITUTE, OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE
NAKAJIMA
COMMENT OS Homo sapiens (human)
PN WO 02052005-A/13
PD 04-JUL-2002
PF 20-DEC-2001 WO 2001JP011217
PR 22-DEC-2000 JP 00P 389742
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/12, C07K14/47
CC Novel gene and protein encoded thereby
FH Key location/Qualifiers
FT CDS location/Qualifiers
(184)..(2178).
1..4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 9.69e-150 Length: 4790
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: Gaps: 6

US-10-029-345A-109_COPY_1_302 (1-302) x BD171157 (1-4790)

QY 1 MetAlaHisGluMetCileGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 184 ATGGCCCATGAGATGATTGGAACCAATTGTTACTGAGAGTTGGTGGCTCTGGCGAA 243

QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
DB 244 AGTGGAAAGGAAAGAGCTGCTTAATTGATAGCCGGCCATTTGTGGAAATCAATACATCC 303

QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIleLeuMetCysValArgLeuGln 60
DB 304 CACATTTTGGAGCATTAAATATCACTCCCAAGCTTATGAAGCGAAGTTGCAACAG 363

QY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleValIleValIleAsp 80
DB 364 GACAAAGTTTAATTAACAGAGCTCATCAGCATTCAGCGAAACATTAAGTTGACATTGAT 423

QY 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIleValIle 100
DB 424 TGCAGTCAGAAAGTTGATGATTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 483

QY 101 AspCysPheLeuThrValIleLeuGlyIleLeuGluIleValIleValIleValIle 120
DB 484 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAGAAAGAGCTTCACTGTTCACTG 543

QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuGlyIleValSer 140

DB 544 CTTCGAGTGGGGTTTCTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGAGGAAATCC 603
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 604 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTGCAACATTTGGCCAAACC 663
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGlyLeu 180
DB 664 CGAATTCCTCCCAATCTTATCTTGGCTGCCAGCGAGATGCTCTCAACAAAGAGCTGATG 723
QY 181 GlnGlnAsnGlyIleGlyTyrValIleAsnAlaSerTyrThrCysProIleProAspPhe 200
DB 724 CAGCAAAATGGAGATTGTTATGTTAAATGCCAGCAAAATACCTGCAAAAGCTGACTTT 783
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIleVal 220
DB 784 ATCCCGAGTCTCATTTCTCGCGTGGCTGTGATGACAGAGCTTTGTGGAATAATTTTG 843
QY 221 ProThrLeuAspIleSerValAspPheIleGluValAlaValAsnGlyCysVal 240
DB 844 CCGTGGTTGGACAAATCAGTATGATTTCAATTGAGAAAGCAAAAGCTTCAATGGATGTT 903
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
DB 904 CTAGTCACTGTTTACCTGGATCTCCGCTCCGCCACATCCCTATCCCTACATCATG 963
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluIleValArgProThr 280
DB 964 AAGAGAGTGAACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAAGACTACT 1023
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluIleValIleVal 300
DB 1024 AATATCTCCAAACTTCAATTTTCTGGGCCCAATCTGACTATGAGAAAGATTAAAGAAC 1083
QY 301 GlnThr 302
DB 1084 CAGACT 1089

Search completed: June 22, 2004, 02:43:22
Job time : 3555.67 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 20:53:18 ; Search time 2262.97 Seconds
(without alignments)
3985.196 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMIGTQVTERIVALL.....PNFPIGLLDYREKIKNGT 302

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgr2_1/USPTO.spool/US10029345/runat_21062004.164146.29039/app_query.fasta_1.1294
-DB=EST -OPMT=fastap -SUFFIX=ext -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloums62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=ptco -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US10029345_0CGC_1_1_3931_@runat_21062004.164146.29039 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST :
1: em_estda :
2: em_esthum :
3: em_estin :
4: em_estnu :
5: em_estov :
6: em_estcp1 :
7: em_estro :
8: em_hic :
9: gb_est1 :
10: gb_est2 :
11: gb_hic :
12: gb_est3 :
13: gb_est4 :
14: gb_est5 :
15: em_estfun :
16: em_estom :
17: em_g88_hum :
18: em_g88_hiv :
19: em_g88_pln :
20: em_g88_vic :
21: em_g88_fun :
22: em_g88_mam :
23: em_g88_mus :
24: em_g88_pro :
25: em_g88_rtd :
26: em_g88_pbg :
27: em_g88_vrl :
28: gb_g88f :

29: gb_g882 :*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1369.5	87.6	898	13 BQ721265	BQ721265 AGENCOURT
2	1145.5	73.2	836	12 B1821804	B1821804 603035883
3	1099	70.3	715	14 CF727177	CF727177 UI-M-HBO-
4	1086.5	69.5	682	12 B1816954	B1816954 imagegc.1
5	1081	69.1	656	14 CF532917	CF532917 UI-M-GHO-
6	1076	68.8	639	14 CF742387	CF742387 UI-M-HBO-
7	1069	68.4	769	13 BU704078	BU704078 UI-M-FHO-
8	1006	64.3	650	10 BF472046	BF472046 UI-M-BH3-
9	989	63.2	2207	11 BC038231	BC038231 Homo sapi
10	978.5	62.6	924	14 CA474739	CA474739 AGENCOURT
11	959	61.3	862	13 BQ770036	BQ770036 UI-M-F10-
12	936	59.8	602	10 AM847425	AM847425 RCI-CT020
13	931	59.5	595	10 AM847426	AM847426 RCI-CT020
14	891	57.0	3325	11 AK035652	AK035652 Mus muscu
15	872	55.8	792	13 BU946569	BU946569 703769910
16	853	54.5	946	12 B1657528	B1657528 603281866
17	823.5	52.7	516	29 AY413655	AY413655 Homo sapi
18	804.5	51.4	516	29 AY413657	AY413657 Mus muscu
19	772.5	49.4	683	12 BM942528	BM942528 UI-M-CGDP
20	770	49.2	619	29 CG583869	CG583869 OST226476
21	769	49.2	579	10 BF582366	BF582366 602094304
22	762	48.7	512	10 BF815601	BF815601 MR2-CT012
23	725	46.4	1060	12 BM546940	BM546940 AGENCOURT
24	703	44.9	421	10 BE946884	BE946884 UI-M-BH3-
25	697.5	44.6	1180	12 BM908762	BM908762 AGENCOURT
26	695	44.4	404	10 BB816652	BB816652 BB816652
27	693	44.3	463	10 BG058779	BG058779 naef10f04.
28	675.5	43.2	412	9 AA878984	AA878984 v001d07.r
29	675	43.2	903	12 BG482429	BG482429 602526595
30	665	42.5	442	14 CB606149	CB606149 AMGNNUC:N
31	659.5	42.2	881	14 CF266082	CF266082 AGENCOURT
32	645	41.2	402	29 CG499793	CG499793 OST42180
33	637	40.7	463	10 BF809325	BF809325 OV0-CT020
34	635	40.6	667	14 CA394418	CA394418 C851C09.Y
35	630.5	40.3	801	14 CK130134	CK130134 AGENCOURT
36	615.5	39.4	685	14 CB165612	CB165612 BEQ603161
37	608	38.9	459	10 BF566061	BF566061 UI-R-BT1-
38	597.5	38.2	403	10 BE004625	BE004625 CM2-BN010
39	596.5	38.1	1148	12 BM806608	BM806608 AGENCOURT
40	586	37.5	494	13 BQ318152	BQ318152 RCI-CT04
41	586	37.5	494	13 BQ318155	BQ318155 RCI-CT046
42	582	37.2	861	14 CA325142	CA325142 UI-M-FYO-
43	577	36.9	635	10 BE317720	BE317720 601229866
44	571	36.5	1043	12 BM563030	BM563030 AGENCOURT
45	569	36.4	337	14 CB606296	CB606296 AMGNNUC:S

ALIGNMENTS

RESULT 1
BQ721265
LOCUS BQ721265 898 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT_8294508 lukeki_synpachetic_trunk Homo sapiens cDNA clone
IMAG:6194455 5', mRNA sequence.
ACCESSION BQ721265
VERSION BQ721265.1 GI:21860162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 898)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM11599 row: 1 column: 08
High quality sequence stop: 669.

FEATURES
source
1. .898
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6194455"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult" 16 yr"
/lab_host="DH10B"
/clone_lib="lupski_sympathetic_trunk"
/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCGACGCGCG-3' and
5'-GACTAGTTCAGTACGAGGAGCGCCGCTT(15)-3'. Site selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) / available through Life
Technologies."

ORIGIN

Alignment Scores:
Pred. No.: 1,786-152 Length: 898
Score: 1369.50 Matches: 279
Percent Similarity: 95.27% Conservative: 3
Best Local Similarity: 94.26% Mismatches: 12
Query Match: 87.56% Indels: 4
DB: 13 Gaps: 1

US-10-029-345a-109_COPY_1_302 (1-302) x BQ721265 (1-898)

```
QY 4 GtMetlleglYThGlnlEValThrgluArgLeuValAlaLeuLeuGluSerGlyThr 23
DB 14 GAGATGATGGAACTCAATTTCTTCTAGAGGTTGGTGGCTCTGCTGGAAGTGGACG 73
QY 24 GtLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 43
DB 74 GAAAAAGTGTCTAATGATGATGCGGCGCATTTGGAAATACATATCCACATTTTG 133
QY 44 GtAlaIleAsnIleAsnCyserIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 63
DB 134 GAAAGCATTAATATCACTGCTCCCAAGCTTTTGAAGCGAAGGTTGCAACGAAAGTG 193
QY 64 LeuIleThrgluLeuIleGlnHisSerAlaIleValIleValIleValIleVal 83
DB 194 TTAATTTACAGAGCTCATCAGCATTCAGCAACATTAAGTTGATGATTTGAGTCA 253
QY 84 LysValIleValIleValIleValIleValIleValIleValIleValIleVal 103
DB 254 AAGGTGTAGTTTACATCAAGCTCCCAAGATTTTCTCTCTCTTCAAGCTGTTT 313
QY 104 LeuThrValLeuLeuGluGlyIleGluGluSerPheAsnSerValIleLeuLeuAla 123
DB 314 CTCACGTGATCTTGGGGTAACTGGAGAAAGCTTCAACTGTTCACTGCTTGGAGGT 373
QY 124 GlyPheAlaGluPheSerArgCysPheProGlyLeuGluGluGlyIleValIleVal 143
```

```
DB 374 GGGTTTGCTGAGTCTCTCGTGTGTTTCCCTGGCCTCTGTGAAGAAATCCACTAGTC 433
QY 144 ProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThrArgIleLeu 163
DB 434 CCTACTGCTGATTTCTGACCTCTGCTTACTGTTTCCACATTTGGCCCAACCCGAATCTT 493
QY 164 ProAsnLeuIleGluGlyCysGlnArgAspValLeuAsnIleGluLeuIleGlnIleAsn 183
DB 494 CCCATCTTTTATCTTGGCTGCGCAGAGATGCTTCAACAAGAGCTGATGACGAGCAAT 553
QY 184 GlyIleGlyIleValIleAsnAlaSerIleThrCysProIleProAspPheIleProGlu 203
DB 554 GGGATGTTGTTTGTGTTTAAATGCGAATACCTGCTCAAGGCTGACTTATCCCGAG 613
QY 204 SerHisPheLeuArgValProValAsnAspSerPheCysGluIleValIleProIlePleu 223
DB 614 TCTCATTTTCTGCGGTGCGCTGGAATGACAGCTTTTGAGAAATTTTGGCCGTGTTG 672
QY 224 AspIleSerValAspPheIleGluValAlaValAlaSerAsnGlyCysValLeuValHis 243
DB 673 GACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTCCCAATGATGTTCTTAGTGAC 732
QY 244 CysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleValIleValIleVal 263
DB 733 TGTTTACTGAGATCTCCCGCTCGGC-ACCATGCTATGCTTACATCAAGAAAGATG 791
QY 264 AspMetSer-LeuAspGluAlaIleValIleValIleValIleValIleValIleVal 283
DB 792 GACATGCTTTTGAAGAACTTACAGATTTTGGAAAGAAAGCAAGCACTACTATATCTCC 851
QY 283 cAsnPheAsn---PheLeuGlyGlnLeuLeuAspIleValIleValIleValIleVal 297
DB 852 AAAAATTCAATTTTCTTGGGCCAAGCTCCCTGCAATTATGAAAAA 897
```

RESULT 2
BI821804
LOCUS 836 bp mRNA linear EST 04-OCT-2001
DEFINITION 603035883F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176724 5',
mRNA sequence.
ACCESSION BI821804
VERSION BI821804.1 GI:15933354
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 836)
TITLE NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strusberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM11440 row: f column: 21
High quality sequence stop: 805.

FEATURES
source
1. .836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5176724"
/lab_host="DH10B"
/clone_lib="NIH MGC 115"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is

oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:
9.27e-126	836	244	1
Score:	1145.50		
Percent Similarity:	96.08%	Mismatches:	8
Best Local Similarity:	95.69%	Indels:	7
Query Match:	73.24%	Gaps:	0
DB:	12		

US-10-029-345A-109_COPY_1_302 (1-302) x B1821804 (1-836)

```

QY 50 CysSerIysLeuMetLysArgArgLeuGlnGlnAspValLeuIleThrGlnLeuIle
DB 2 TGCCTCCAGGCTTATGAAAGCGAGGTGCAACAGACAGAGCTTATACAGAGCTCATC 61
QY 70 GlnHisSerAlaIysHisIysValAspIleAspCysSerGlnIysValValIlyrAsp 89
DB 62 CAGCATTCAGCGAAACATAAGTTGACATTGA-TGCAGTCAGAAAGGTTGAGTTACGAT 120
QY 90 GlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGly 109
DB 121 CAAGAGCCCCAAGATGT-GCCTCTCTCTCTCAGACAGCTTTCTCAGCTGACTTCGGGT 179
QY 110 LysLeuGlnLysSerPheAsnSerValHisLeuLeuAlaGlyIlePheAlaGlnPheSer 129
DB 180 AAATCGAGAAAGAGCTTCAACTGTCACCTCTGACAGTGGTGGTGGTGGTGGTCTCT 239
QY 130 ArgCysPheProGlyLeuCysGlnGlyLysSerThrLeuValProThrCysIleSerGln 149
DB 240 CGTAGTTTCCCTGGCTCTGTAAGAAATCCACTACTACTCTGCTGCTGCTGCTGCTG 299
QY 150 ProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuIlyrLeuGly 169
DB 300 CTTTGGCTACTCTTGGCCAACTGGGGCCAAACCGGATCTTCCCAATCTTATCTTGGC 359
QY 170 CysGlnArgAspValLeuAsnLysGlnLeuIleGlnGlnAsnGlyIleGlyIlyrValLeu 189
DB 360 TGCAGAGAGATGCTCTCAACAGAGAGCTGATGAGAGATGGATGGATGGATGGATG 419
QY 190 AsnAlaSerIyThrCysProLysProAspPheIleProGlnSerHisPheLeuArgVal 209
DB 420 AATGCCAGCAATACCTCTCCAAAGCCGACTTATCCCGAGTCTCATTTCCGCGGTG 479
QY 210 ProValAsnAspSerPheCysGlnLysIleLeuProThrLeuAspIlyrValAspPhe 229
DB 480 CCGTGTGATACAGCTTCTGTGAGAAATTTTCCCGTGGTGGACAAATCAGTAGATTTTC 539
QY 230 IlleGlnLysAlaIysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSer 249
DB 540 ATTAGAGAAACAAAGAGCTCCAAAGATGATGT-CTAGTGCAGCTTTAGCTGGAGATCTCC 588
QY 250 ArgSerAlaThrIleAlaIleAlaIyTrIleMetLysArgMetAspMetSerLeuAspGlu 269
DB 599 CGCTCCGC-ACCATCGCTATCGC-TACATCATGAAAGAGATGACATGCTTTTNGATGA 656
QY 270 AlaIyTrArgPheValLysGlu-LysArgProThrIleSerProAsnPheAsnPheLeu-G 289
DB 657 GCTTACAGATTTGTGACAGAAAGAGACCTACTATATCTCCAAACTTCAATTTTTCGGG 716
QY 289 LysGlnLeuAspIyThrGlnLysLysIleLysAsnGlnIlyr 302
DB 717 GCCAGCTCTCGAGCTATGAGAAAGATTAAGAACCAAGACT 757

```

RESULT 3
CF727177 715 bp mRNA linear EST 09-OCT-2003

DEFINITION UI-M-HB0-ckj-3-09-0-UI.r1 NIH BMAP_HB0 Mus musculus cDNA clone
IMAGE:30548096 5', mRNA sequence.
ACCESSION CF727177
VERSION CF727177.1 GI:37601345
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS NIH-MGC http://mgi.mgi.nih.gov/.
TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

seq primer: pyx-5.
Location/Qualifiers
1.715
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30548096"
/issue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1b="NIH BMAP_HB0"
/note="Organ: Eye; Vector: pyx-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTAATGAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

FEATURES
source

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:
2.58e-120	715	206	1
Score:	1099.00		
Percent Similarity:	98.15%	Mismatches:	4
Best Local Similarity:	95.37%	Indels:	0
Query Match:	70.27%	Gaps:	0
DB:	14		

US-10-029-345A-109_COPY_1_302 (1-302) x CF727177 (1-715)

```

QY 87 ValIyTrArgGlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrVal 106
DB 2 GTTATGATCAAGAGTCCCAAGATGTGGTCTCTGTCTGTCAGAGCTTTTCCATCTGA 61
QY 107 LeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeuLeuAlaGlyIlePheAla 126
DB 62 CTTCTGGGTAACAGAGAGAGAGAGCTTCAACTGTCACCTGCTTCCAGGTGGCTTGGCT 121
QY 127 GluPheSerArgCysPheProGlyLeuCysGlnGlyLysSerThrLeuValProThrCys 146
DB 122 GAGTCTCTCGTGTGTTCCCTGGCTCTGTGAAAGAAATGCCACTGACTGCTGCTGCTG 181

```

```
QY 147 IISerGlnProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnIleu 166
|||
Db 182 AATCTCAGCCTTCTGCTTACCTGTGCGAAACATGGGCCAAGCTCGAATCTTCCCAATCTC 241
QY 167 TTTLeuGlyCysGlnArgAspValLeuAsnIleGlyLeuIleGlnIleAsnGlyIleGly 186
242 TATCTTGGGTGCGCAGGAGATGCTCTCAACAAGAGACCTGAGCAACAGATGGATTGGC 301
QY 187 TTTValLeuAsnAlaSerTyrThrCysProIysProAspPheIleProGluSerHisPhe 206
302 TATGCTTTAAATGCGCAGCAATATCTGTCCAAAGCCTGATCTTCAATCCTGAATCTCACTTC 361
QY 207 LeuArgValProValAsnAspSerPheCysGlnIleuIleLeuProTyrLeuAspIleuSer 226
362 CTGCGAGTGCTGTGAATGACAGCTTTTGTGAGAAATCTTACCATGTGTGGACAGATCT 421
QY 227 ValAspPheIleGlyValAlaValAlaSerAsnGlyCysValLeuValHisCysLeuAla 246
422 GTGATTTTCATTGAGAAAGCAAAAGCCTCCAAATGGCTGTGTGCTTATCCACTGCTTAGCT 481
QY 247 G1YIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetIysArgMetAspMetSer 266
482 GGGATCTCTGCTCGGCACATATGCTATTGCTTACATCATGAGAGAGATGACATGTCT 541
QY 267 LeuAspGlnAlaTyrArgPheValIleGlyIleArgProThrIleSerProAsnPheAsn 286
542 CTAGATAGAGCTTACATATTTGTGAAGAAAGAAAGAAAGCTTACTATCTCGAATTTTAT 601
QY 287 PheLeuGlyIleLeuLeuAspTyrGlyIleValIleLeuAsnGlnThr 302
602 TTTATGGCCACTCATGACATGATGAGAGACATTAATTAACAGACT 649
```

RESULT 4

```
BI816954/ 682 bp mRNA linear EST 10-OCT-2001
LOCUS imagec.10.2000/s12410bdf41.x1 Soares_NPBMC Homo sapiens cDNA
DEFINITION clone IMAGE:4140798 3', mRNA sequence.
ACCESSION BI816954
VERSION BI816954.1 GI:15911639
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and
AUTHORS Prange, C.K.
TITLE The I.M.A.G.E. Consortium quality control effort: clone
JOURNAL resequencing for verification
COMMENT Unpublished (2001)
Other ESTs: BG058779
Contact: Prange CK
The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA
Email: help@image.llnl.gov
This read has been verified (found to hit its original self in the
correct orientation), as part of the I.M.A.G.E. Consortium quality
control effort. High quality sequence is defined as having 100 or
more base pairs with a phred quality value of 20 or greater, where
a sliding window of 4 base pairs with a phred quality value of 15
or greater marks the beginning and end of the sequence. For
information on obtaining this clone, please contact
info@image.llnl.gov. effort.
Plate: LLM9388 row: k column: 7
Seq primer: -21m13
High quality sequence stop: 682.
Location/Qualifiers
```

FEATURES

```
1..682
/mol_type="mRNA"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4140798"
```

```
/tissue_type="lymphocyte"
/lab_host="pH10b (phage-resistant)"
/clone_id="Soares_NPBMC"
/note="Organ: blood; Vector: pT73D-Pac; Site 1: NotI;
Site 2: EcoRI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTTCACATCTGAAGTGGGAGCGCGCGGGTGTGTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
is normalized, constructed in the laboratory of M. Bento
Soares (University of Iowa)."
```

ALIGNMENT SCORES:

Pred. No.:	7,44e-119	Length:	682
Score:	1086.50	Matches:	214
Best Local Similarity:	96.89%	Conservative:	4
Query Match:	95.11%	Mismatches:	6
DB:	69.47%	Indels:	2
	12	Gaps:	1

US-10-029-345A-109_COPY_1_302 (1-302) x BI816954 (1-682)

```
QY 49 AsnCysSerIysLeuMetIysArgArgLeuGlnIleAspIleValIleuIleThrGluLeu 68
|||
Db 673 AACTGTCACAGCTA--TAAAGCAAGTGCGCAACAGACAAAGTGT-ATTCAGAGCTC 618
QY 69 IISerIleSerAlaIysHisIysValAlaPheIleAspCysSerGlnIleValIleuIle 88
617 ATCCAGATTCACCGAAACATTAAGCTTGACATTTGATGACATGAGAGGTGTGATGTAC 558
QY 89 AspGlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrValLeuLeu 108
557 GATCAAGAGCTCCCAAAATGTCTCTCTCTCTTCAAGACTGTTTCTCACTGTAATTCTG 498
QY 109 G1YIleGlnIleuIysSerPheAsnSerValHisLeuAlaIleGlyIlePheAlaGluPhe 128
497 GGTAACTGAGAGAGAGCTTCAACTGTGTCACCTGTCGAGGTGGAGGTGCTGAGCTTC 438
QY 129 SerArgCysPheProGlyLeuCysGlnIleuIysSerThrLeuValProThrCysIleSer 148
437 TCTCGTGTGTTTCCCTGCGCTCTGTGAAGAAATCACTCAATGTCCTACCTGCAATTCT 378
QY 149 GlnProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrIleu 168
377 CAGCTTGCTTACCTGTGCGCAACATTTGGGCCAACCCGAATCTTCCCAATCTTATCTT 318
QY 169 G1YCysGlnArgAspValLeuAsnIleuIleGlnIleGlnIleAsnGlyIleGlyTyrVal 188
317 GGCTGCCAGCGAGATGCTCTCAACAAGAGCTGATGACAGCAAGATGGAGTTGGATTATGT 258
QY 189 LeuAsnAlaSerTyrThrCysProIysProAspPheIleProGluSerHisPheLeuArg 208
257 TTTAAATGCGCAGCAATATCTGTCCAAAGCTGACTTATCCCGAGTCTCATTTCTCGT 198
QY 209 ValProValAsnAspSerPheCysGlnIleuIleuProThrLeuAspIleuSerValAsp 228
197 GTGCTGTGATGACAGCTTTTGTGAAGAAATTTTGGCGGTGTGGAACAATATGATGAT 138
QY 229 PheIleGlyValAlaValAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIle 248
137 TTTCAITGAGAAAGCAAAAGCTCCCAATGATGTGTTCTTACTGACATGTTAGCTGGAGTC 78
QY 249 SerArgSerAlaThrIleAlaIleAlaTyrIleMetIysArgMetAspMetSerLeuAsp 268
77 TCCGCTCCGCGACCATGCTTATGCTTACATATGATAGAGATGACATGTCTTATGAT 18
QY 269 GlnAlaTyrArgPhe 273
Db 17 GAACTTACAGATT 3
```

RESULT 5

CF532917 656 bp mRNA linear EST 12-SEP-2003
DEFINITION UI-M-GH0-cgw-n-02-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
IMAGE:30357529 5', mRNA sequence.
ACCESSION CF532917
VERSION CF532917.1 GI:34584885
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 656)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES
source
1. 656
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30357529"
/dev_stage="1, 5, and 15 days newborn"
/tissue_type="whole brain"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH_BMAP_GH0"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGACTGAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

Db 123 GCAGAGTTGCACACAGACAAAGATTAATTAACAGAACTTATCAACATCTCGCAAGCA 182
Qy 75 blyValaPp1laPcysSerGlnyValaValaValaYrAspGlnSerGlnaPpva 95
Db 183 TAAAGTTGACATTAAGTCAATGAGAGAGGAGGAGTTAAGATCAAGATTCACAAAGATGT 242
Qy 95 lAlaSerLeuSerSerAPcysPheLeuThrValLeuLeuGlyLeuLeuGlyLeuSerPh 115
Db 243 TGGTTCTGTGTCTGACATGCTGCTTCTCACTGACTGTTCTGGGTAACAGAGAGACTT 302
Qy 115 eaenSerValhLeuLeuAaGlyGlyPhealagluPhealagluPheSerArgCysPheProGlyLe 135
Db 303 CAACCTGTGTCACCTGCTGTCAGAGTGGCTTGTGAGTTCTCGTTGTTCCCTGGCCT 362
Qy 135 uCySgluGlyLysSerThrLeuValProThrCySgluSerGlnProCysLeuProValAl 155
Db 363 CTGTGAAGGAATGCACTCTACTCTCCCTCACTCAATCTCAGCTTGTACTTACTTGTGC 422
Qy 155 aAenlEGlyProThArGllLeuProaenLeuTyLeuGlyCySgluArgaPpValle 175
Db 423 GAACATTTGGCCCACTGAAATTTCTCCATCTCTATCTTGGCTGCCAGAGATGCTCT 482
Qy 175 uAenlGlyLeuLeuGlnGlnaGlylleglTyYrValLeuAaPp1aSerTyThrCy 195
Db 483 CACAAAGACCTATGACACAGATGGAATGGCTATGTTAAAGCCAGCAATACCTG 542
Qy 195 sProlYsPProAPpHe1leProGluSerh1sPheLeuAaGValProValaPp1aSerPh 215
Db 543 TCCAAAGCCCTGACCTTATCTCAATCTCACTTCCGCGAGTCCGTGTGAATACAGCTT 602
Qy 215 eCySgluValleLeuProThPheuAaPp1aSerValaPpHe 229
Db 603 TTGTGAAGAAATCTACCATGTTGACAGATGTCATGAGATTTC 645
RESULT 6
CF742387 639 bp mRNA linear EST 10-OCT-2003
DEFINITION UI-M-HB0-cl1-c-20-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
IMAGE:30619363 5', mRNA sequence.
ACCESSION CF742387
VERSION CF742387.1 GI:37638726
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 639)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES
source
1. 639
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30619363"
/tissue_type="whole eye"
/dev_stage="embryo 12.5, 13.5, 14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH_BMAP_HB0"


```

Db      126 CTGGCCCTGTGAGGAAGTCCACTCTGCTCCTACTGATATCTCAGCTTGCTTA 185
Qy      153 ProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrlleuGlyCyseGlnArg 172
Db      186 CCGTGGCGGAACATGGGCCCAACTCGAATCTCCCATCTCTATCTTGCTGCCACGGA 245
Qy      173 AspValIleuAsnIleGlyLeuIleGlnIleAsnGlyIleGlyTyrlleuAsnAlaSer 192
Db      246 GATGCTCTCAACAGACGCTGATGCAACAGATGGATGGCTATGTGTTAATGCCAGC 305
Qy      193 TyrlleuGlyProAspPheIleProGluSerHisIleLeuArgValProValAsn 212
Db      306 AATACCTGCTCAAGCCTGACTTCACTGATCTGATCTCTCCGAGGCGCTGTAAT 365
Qy      213 AspSerPheGlyGlyIleLeuProThrLeuAspIleSerValAlaPheIleGlyLeu 232
Db      366 GACAGCTTTGTGAGAAATCTCAACATGCTGGCAAGCTGTGAGATTCATTGAGAAA 425
Qy      233 AlaIleAlaSerAsnGlyCyseValIleuValHisCyseLeuAlaGlyIleSerArgSerAla 252
Db      426 GCMAAGCCTCCCAATGCTGCTGCTTATCTCAGCTTACCTGGATCTCTCCGCTCCGC 485
Qy      253 ThrIleAlaIleAlaTyrlleuIleGlyArgMetAspMetSerIleuAspGluAlaTyrlleu 272
Db      486 ACTATGCTATGCTGCTCATCATGAGAGATGAGATGCTCTGATGAGGCTTACAGA 545
Qy      273 PheValIleuGlyIleGlyProThrIleSerProAsnIleAsnIleGlyIleLeu 292
Db      546 TTGTGTGAAGAAAGAAAGAACTACTATCTCCGAATTTTATGAGGCCCAACTCATG 605
Qy      293 AspTyrGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIle 302
Db      606 GACTATGAGAGAGAGATTAATACCAACT 635

```

ORIGIN

/clone="UI-M-BH3-awu-d-10-0-UI"
 /dev stage="27-32 days"
 /lab host="DH10B (Life Technologies)"
 /clone lib="NIH_BMAP_M_S4"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The
 NIH_BMAP_M_S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH_BMAP_M_S4,
 NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
 NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
 (NIH_BMAP_M_S4) was constructed as follows: PCR amplified
 cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
 NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH_BMAP_M_S4 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)"

```

RESULT 8      BF472046      650 bp      mRNA      linear      EST_04-DEC-2000
LOCUS        BF472046
DEFINITION   UI-M-BH3-awu-d-10-0-UI. r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
VERSION      BF472046.1 GI:11541229
KEYWORDS     EST.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus

```

```

REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE         Bonaldo, M.F., Lennon, G. and Soares, M.B.
              Normalization and subtraction: two approaches to facilitate gene
              discovery

```

```

JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
PUBMED       8889548
COMMENT      Contact: Chin, H
              National Institute of Mental Health
              6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
              20892-9643, USA
              Tel: 301 443 1706
              Fax: 301 443 9890
              Email: mestr@mail.nih.gov

```

CDNA library Preparation: M.B. Soares Lab Clone distribution:
 Researchers may obtain BMAP cDNA clones from RESARCH GENETICS. It
 should be noted that Bento Soares is generating a small number of
 additional specialized non-redundant arrays of BMAP cDNAs whose
 availability will be considered under appropriate and limited
 collaborative arrangements
 Seq primer: M13 Reverse.

```

FEATURES
Source       Location/Qualifiers

```

```

1..650
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

```

```

Alignment Scores:
Pred. No.:      2,86e-109
Score:          1006.00
Percent Similarity: 97.21%
Best Local Similarity: 95.35%
Query Match:    64.32%
DB:             10
Gaps:           0

```

US-10-029-345A-109_COPY_1_302 (1-302) x BF472046 (1-650)

```

Qy      5 MetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGluSerGlyThrGlu 24
Db      3 ATATATTGGAACCTCAATTTGTTACTGAGAGCTTGCTGCTGCAAGATGGAACGGA 62
Qy      25 LysValIleuLeuIleAspSerArgProPheValGlyTyrAsnThrSerHisIleGlu 44
Db      63 AATATG-CTGCTAATTGATATGCGACCATTTGTGGAATACAAATGCTCTCAATTTGGAA 121
Qy      45 AlaIleAsnIleAsnCyseSerIleuMetIleValArgIleuGlnIleAspIleValIleu 64
Db      122 GCCATTAAATATCAACCTCCCAACATGAGACCGAAGCTTGCAACAGCAAGAAATTA 181
Qy      65 IleThrGluLeuIleGlnHisSerAlaLysHisIleValAspIleAspCyseSerGlnLys 84
Db      182 ATTACGAACCTTATCCAACTTTGCAAGACATTAAGCTTACATTCGCAATCCAGAG 241
Qy      85 ValValIleValTyrAspGlnSerSerGlnAspValAlaSerIleuSerAspCysePheLeu 104
Db      242 GTGTGATTATGATCAAGATTCCCAAGATGTTGCTCTGTCGTCAGACTGCTTTCTC 301
Qy      105 ThrValIleuLeuGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIle 124
Db      302 ACTGACTCTTGCTGATTAATGAGAGAGACTTCACCTGTCACCTGTCGAGGTGGC 361
Qy      125 PheAlaGluPheSerArgCysePheProGlyLeuCyseGluGlyLysSerThrIleuValPro 144
Db      362 TTTCGTGAGTTCTCTGTTGTTCTCTGCTGCTCTGTCGAGAGAAA-TCCACTCTAGTCCCT 420
Qy      145 ThrCysIleSerGlnProCyseLeuProValAlaAsnIleGlyProThrArgIleuPro 164
Db      421 ACTGTGATATCTCAGCTTGCTTACTGTTGGAAATTCATGGGCAATTCGATTTCTCC 480

```

OY		165	AstleuryTtlenugjyCysglnaragapValleuanenlysgluleullleglngInaengly	184
Dd		481	AATCTCTATTCTGGCTGCACAGCAGATGTCTCCAAACAAGGACTGAAGCAACGAATGCG	540
OY		185	IlleglyTyrtAlleunanaIsertyrThCySpPolysProAepPheileProgiUser	204
Dd		541	ATTGGCTATVTGTTTAATACCAGCAATACCTGTCCAAGGCTGACTTCATACTGAATCT	600
OY		205	HlsphelenuaryalProvalAsnaSpssePhcysgluysslle	219
Dd		601	CAC TTCCTCGAGTGCC TGTGATATACAC -TTTTGTGAGAAATTC	644
RESULT 9 BC038231 LOCUS		BC038231	2207 bp mRNA linear HTC 01-OCT-2002	
DEFINITION			Homo sapiens, Similar to dual specificity phosphatase 8, clone IMAGE:5547764, mRNA.	
ACCESSION		BC038231		
VERSION		BC038231.1	GI:23398534	
KEYWORDS		HTC.		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Euteleia; Primates; Catarrhini; Homnidae; Homo.		
TITLE		1 (bases 1 to 2207)		
JOURNAL		Strausberg R. Direct Submision Submitted (30-SEP-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK		NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT		Contact: MGC help desk		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: <http://image.inl.gov>
 Series: IRK Plate: 79 Row: A Column: 12
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4758211
 This clone has the following problem: frame shifted.
rock@rockefeller.edu

```

source
1. .2207
/organism="Homo sapiens"
/mol_type="mrna"
/bi_xref="taxon:9606"
/clone="MAGE:5547764"
/issue_type="Skin, melanotic melanoma."
/lab_id="NIR MGC_72"
/lab_host="DH10B"
/notes="Vector: pCMV-Sport6"

```

```
Alignment Scores: 2.31e-106
Pred. No.:
Score: 989.00
```

Percent Similarity:	81.40%	Conservative:	55
Best Local Similarity:	63.12%	Mismatches:	54
Query Match:	63.24%	Indels:	2
DB:	11	Gaps:	2

US-10-029-345A-109_COPY_1_302 (1-302) x BC038231 (1-2207)

[illegible]

LOCUS CA474739 924 bp mRNA linear EST 12-NOV-2002
 DEFINITION AGENCOURT 10700573 NCI_CGAP_ZKId1 Danio rerio cDNA clone
 IMAGE:6797004 5', mRNA sequence.
 ACCESSION CA474739
 VERSION CA474739.1 GI:24931091
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 924)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Leonard I. Zon, M.D.
 CDNA Library Preparation: Invitrogen Corp.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM14309 row: j column: 11
 High quality sequence start: 24
 High quality sequence stop: 708.
 Location/Qualifiers
 1..924
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:6797004"
 /lab_host="DH10B (T1-resistant)"
 /clone_1db="NCI CGAP ZKId1"
 /note="Organ: kidney; Vector: PCMV-SPORE.1; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Constructed by J. Wang (Research Genetics, Invitrogen Corp) from tissue donated by L. Zon (Harvard University). Note: this is a NCI_CGAP Library."
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,89e-106 Length: 924
 Score: 978.50 Matches: 203
 Percent Similarity: 86.08% Conservative: 32
 Best Local Similarity: 74.36% Mismatches: 33
 Query Match: 62.56% Indels: 6
 Gaps: 1
 DB: 14
 US-10-029-345a-109_COPY_1_302 (1-302) x CA474739 (1-924)
 QY 13 GlnAgtgLeuValAlaLeuLeuGluSerGlyThrcjLutyrValleuLeuLeuLeuAapSerArg 32
 Db ::
 118 CAGGCTTGGTGGCGCTGCGAGGCGCGGTGAGCGGCTGCTCTCATCTGACGAGG 177
 QY 33 ProPheValGluTyrAaThrSerHisIleLeuGluAlaIleAaniLeaAncySerLys 52
 Db ::
 178 CCCTGCTGGAGTTCAACTCTGCAATCTGAGGCGCGGAACATCAACTGCTCCAG 237
 QY 53 LeuMetLysArgArgLeuGlnGlnAapLysValleuLeuLeuThrGluLeuLeuLeuLeuSer 72
 Db ::
 238 CTGATGAAGAGAGACAGCAGCAGACAGAAATCCAGATCCAGCCGAACCTGCGACACTGC 297
 QY 73 AlaLysHisLysValAapLysSerGlnLysValValValTyrAapGlnSerSer 92
 Db ::
 298 GCCAAG 354
 QY 93 GlnAapValAlaSerLeuSerSerAapCyAphLeuThrValleuLeuGluLysLeuGlu 112
 Db ::
 355 GGATGACCCCGCTCTCTCAACCGATGCCCTTCTCAGCGCTGTTGTCAGAGCTGAG 414

QY 113 LysSerPheAasSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArgCyAph 132
 Db ::
 415 AAGAGTTTCCATCATCTCACTGCTCTCAAGGGGTTTCTCGAGTTCTCAGGGCTGTT 474
 QY 133 ProGlyLeuCyAagGluGlyLysSerThrLeuValProThrCyAileSerGlnProCyA 152
 Db ::
 475 CCGGCTGTGTGAGGGGAAGCTGCGCTGCTCCCTGTGTGTCTGCAAGTGTGTG 534
 QY 153 ProValAlaAaniLeuLysProThrArgGlyLeuProAaniLeuLysGlyCyA 172
 Db ::
 535 TCGGTGAGAGCGCGCGCCAGCGGATCTGCAACCTGTACCTGCGGCTCCAGCGG 594
 QY 173 AapValLeuAaniLysGluLeuLeuGlnGlnAaGlyIleGlyTyrValleuAaniA 192
 Db ::
 595 GACGTCTCAACAGAGAGCTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
 QY 193 TyrThrCyAProLysProAapPheLeuProGluSerHisPheLeuArgValProValA 212
 Db ::
 655 AACTCTGCCCCAAGCAGACTTCATCCGAGACACACTTCTGCGGGTCCGAGAAC 714
 QY 213 AapSerPheCyAglLysLysLeuProTyrPheAapLysSerValAapPheLeuLys 232
 Db ::
 715 GACAGCTTCTGCGAAGATCTGCGCTGCTGACCGCTCGTGGAGTTTCATGAGAAA 774
 QY 233 AlaLysAlaSerAaniLysValleuValHisCyALeuAlaGlyLysSerArgSerAla 252
 Db ::
 775 GCCAAGGCCAGTACCCAGAGATTCTGCTTCACTGCTGCGGGAAATCTCCGCTCGCC 834
 QY 253 ThrIleAlaIleAaniLysLysMetLysArgMetAapMetSerLeuAapGluAlaTyr 271
 Db ::
 835 ACATC-GCATCGCTCATCATGAAAGAGATGACATGACACTGGAGTAAACGTAAC 893
 QY 272 ArgPhe-ValLysGluLysArgProThrIle 281
 Db ::
 894 AGGTTGGAGAAAGAGAGAGAGCGCCGACATC 924
 RESULT 11
 B0770036 862 bp mRNA linear EST 26-JUL-2002
 LOCUS B0770036
 DEFINITION UI-M-F10-byc-o-24-0-UI-r1 NIH_BMAP_F10 Mus musculus cDNA clone
 IMAGE:5702255 5', mRNA sequence.
 ACCESSION B0770036
 VERSION B0770036
 KEYWORDS B0770036.1 GI:21978510
 SOURCE EST.
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 1 (bases 1 to 862)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 Seg primer: pYX-5.
 Location/Qualifiers
 1..862
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:5702255"
 /issue_type="whole brain"
 /dev_stage="embryo 12.5dpc"

/lab_host="DH10B (T1 phage resistant)"
 /clone_1ib="NIH BMAP F10"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: Bcor I;
 Site 2: Not I; The library was constructed according
 Ronaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Bcor I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CAGCCACGAC. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemlin Chin, Ph.D.,
 program coordinator."

ALIGNMENT SCORES:

Pred. No.:	1,89e-103	Length:	862
Score:	959.00	Matches:	178
Percent Similarity:	98.40%	Conservative:	6
Best Local Similarity:	95.19%	Mismatches:	3
Query Match:	61.32%	Indels:	0
DB:	13	Gaps:	0

US-10-029-345A-109_COPY_1_302 (1-302) x BQ770036 (1-862)

```

QY 116 AanserValH1eLeuLeuA1aG1yG1phea1aG1upheserArgCysePheProG1yLeu 135
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 8 CACTCTGTCCACTCTGCTTCAGAGTGGCTTGTGAGTCTCTCGTGTGTTCCCGGCTC 67

QY 136 CysG1uG1yLysSerThrLeuValProThrCys11eSerGlnProCysLeuProValA1a 155
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 68 TGTGAGAGAAATCCACTAGTCCCTACCTGCTCAATCTCAATCTGCTTACCTTCTCG 127

QY 156 Aen11eG1yProThrArg11eLeuProAen1eU1yLeuG1yCysG1nArgAspVal1eU 175
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 128 AACATTTGGCCAACTCGAATTTCTCCAAATCTCTATCTGTGGCTGCACGAGATGTCCTC 187

QY 176 Aen1yG1yLeu11eG1n1nAenG1y11eG1yT1yVal11eAenA1aSerT1yThCys 195
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 168 AACAGAGACCTGATGCAACAGATGATGCTATGCTTAAATGCCAGAAATACCTGT 247

QY 196 Pro1ySProAspPhe11eProG1uSerH1sPheLeuArgValProValAsnAspSerPhe 215
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 248 CCAAGCCTGACTTCACTGATCTGATCTCCTGCGAGTGGCTGTGAATGACAGCTTT 307

QY 216 CysG1uYs11eLeuProThrLeuAsp1ySserValAspPhe11eG1uYsAla1yA1a 235
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 308 TGTGAGAAATCTCAATGCTGTGAGCAAGTCTGTGATTCATTGAGAAAGCAAAAGCC 367

QY 236 SerAsnG1yCysVal1eUValH1eCysLeuA1aG1y11eSerArgSerAlaThr11eAla 255
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 368 TCCATGCTGTGCTTATCCACGCTTACCTGGAGATCTTCCTCGCCACCACTATTTGCT 427

QY 256 11eAla1yT11eMet1yArgMetAspSerSerLeuAspG1uA1a1yArgPheVal1yS 275
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 428 ATTGCTTCACTCATGAGAGATGAGATGCTCTAGATGAGGCTTACAGATTTGTGAAA 487

QY 276 G1u1yAsArgProThr11eSerProAsnPha1nPha1nLeuG1y1n1eULeuAspT1yG1u 295
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 488 GAAAAAGACCTACTATATCTCCGAAATTTAATTTATGAGGCCAACTCATGAGACTATGAG 547

QY 296 Lys1yS11e1ySAsnG1nThr 302
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 548 AAGACGATTATTAACCAAGACT 568

```

RESULT 12

AM847425/c AM847425 602 bp mRNA linear EST 19-MAY-2000
 LOCUS RCL-CT0206-270999-021-g01 CT0206 Homo sapiens cDNA, mRNA sequence.

ACCESSION AM847425
 VERSION AM847425.1 GI:7942942
 KEYWORDS EST
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 602)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagaï,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 2020263
 PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=kt2=RCL-CT0206-270
 999-021-g01&t=1999-09-27&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 8
 High quality sequence stop: 602.
 Location/Qualifiers
 1..602
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev stage="Adult"
 /clone_1ib="CT0206"

/note="Organ: colon; Vector: puc18; Site 1: Sma1; Site 2:
 Sma1; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No.
 196,716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."

ORIGIN

Alignment Scores:	5,85e-101	Length:	602
Pred. No.: <td>936.00</td> <td>Matches: <td>190</td> </td>	936.00	Matches: <td>190</td>	190
Score: <td>97.96%</td> <td>Conservative: <td>2</td> </td>	97.96%	Conservative: <td>2</td>	2
Percent Similarity: <td>96.94%</td> <td>Mismatches: <td>4</td> </td>	96.94%	Mismatches: <td>4</td>	4
Best Local Similarity: <td>59.85%</td> <td>Indels: <td>3</td> </td>	59.85%	Indels: <td>3</td>	3
Query Match: <td>10</td> <td>Gaps: <td>0</td> </td>	10	Gaps: <td>0</td>	0

US-10-029-345A-109_COPY_1_302 (1-302) x AM847425 (1-602)

```

QY 58 LeuG1nG1nAsp1yVal1eU11eThrG1uLeu11eG1nH1sSerAla1yS1yVal 77
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 592 GTGCAACAGACCAAGTGTATTAATTCAGAGCTCACGACATTCAGCAAACTAAGGTT 533

QY 78 Asp11eAspCysSerG1n1yVal1yVal1yArgP1nSerSerG1nAspValA1aSer 97
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 532 GACATTGATTGACAGTCCAGAGTTGATTACGATCAAGCTCCCAAGATGTTCTCT-- 475

QY 98 LeuSerSerAspCysPheLeuThrVal1eUleuG1y1yS1eG1uYs1ySserPheAsnSer 117
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 474 CTCCTCTCAGACTGTTTCTCACTACTCTGAGGTAAACGAGAAAGACTTCAACTCT 415

QY 118 ValH1s1eUleuA1aG1yG1yPheA1aG1uPheserArgCysPheProG1yLeuCySg1u 137

```

```

Db 414 GTTCACCTGCTTGCAGGTGGTTGGCTGAGATTCCTCGTGTTCCTGGCCCTCTGAA 355
Qy 138 GtlyvserThrleuValProThrCysIleSerGlnProCysleuProValAlaAnlle 157
Db 354 GGAAAAATCCACTCTAGTCCCTACCTGATTCCTCAGGCTTGCTTACCTGTGGCAACATT 295
Qy 158 GtlyProThrArgIleleuProAnleuTyrlleuGtlyCysGlnArgApValleuAnly 177
Db 294 GGCCCAACCCGAATCTTCCCAATCTTATCTTGCGCCAGGAGATGCTTCAACAG 235
Qy 178 GtlyleuIleGlnGlnAnnglyIleGtlyTyrlValleuAnAlaserTyThrCysProlys 197
Db 234 GACCTATACAGCAGAAATGGATGGTATGTGTAAATCCAGCTATACCTCTCCNAAG 175
Qy 198 ProAspPheIleProGlnSerHisPheleuArgValProValAnAspSerPheCysGlu 217
Db 174 CCGACCTTATCCCGGCTCTCATTTCTGCTGCTGCTGGAATGACAGCTTTGTGAG 115
Qy 218 LysIleleuProThrleuAspIlySerValAspPheIleGtlyValAlaserAn 237
Db 114 AAAATTTTGGCGGTGGACAA-TCAGTAGATTTCATTGAGAAAGCAAGCCTC--CAT 58
Qy 238 GtlyCysValleuValHisCysleuAlaglyIleSerArgSerAlaThr 253
Db 57 GGATGTCTTACTGACCTGTTTACGTGGATCTCCCGCTCCGCCACC 10

RESULT 13
AM847426 595 bp mRNA linear EST 19-MAY-2000
LOCUS R01-CT0206-270999-021-g06 CT0206 Homo sapiens cDNA, mRNA sequence.
DEFINITION AM847426
ACCESSION AM847426
VERSION AM847426.1 GI:7942943
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 595)
Dlae Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Negal,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
JOURNAL CONTACT: Simpson A.J.G.
MEDLINE LABORATORY OF Cancer Genetics
PUBMED LUDWIG Institute for Cancer Research
10737800 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPSP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ctz=RCI-CT0206-270
999-021-g06&ctz=1999-09-27&ct=1)
Seq primer: puc 18 forward
High quality sequence stop: 595.
location/Qualifiers
1. 595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_id="CT0206"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived

```

```

ORIGIN
Alignment Scores:
Pred. No.: 2,27e-100 Length: 595
Score: 931.00 Matches: 189
Percent Similarity: 97.95% Conservative: 2
Beet Local Similarity: 96.92% Mismatches: 4
Query Match: 59.53% Indels: 3
DB: 10 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AM847426 (1-595)
Qy 58 LeuGlnGlnAspIlyValleuIleThrGtlyleuIleGlnHisSerAlaIlyHisIlyVal 77
Db 582 GTGCACAGAGACAAAGTGAATTACAGAGCTCATCCAGCATTCAGCGAAACATAGGTT 523
Qy 78 AspIleAspCysSerGlnIlyValValIlyTyrlAspGlnSerSerGlnAspValAlaser 97
Db 522 GACATTGATTGACGTCAGAAAGTTGATTACGATCAAGCTCCCAAGATGTGCT-- 465
Qy 98 LeuSerSerAspCysPheleuThrValleuLeuGtlyValleuGtlyIlySerPheAsen 117
Db 464 CTCTCTTCAGACTGTTTCTCACTGACTGTTCTGGGTAACCTGAGAAAGACTTCAACTCT 405
Qy 118 ValHisleuLeuAlaglyIlyPheAlagIlyPheAsenArgCysPheProGtlyleuCysGlu 137
Db 404 GTTCACCTGCTGGACAGGTGGTTGTGATGTTCTCGTGTGTTCCCTGGCCCTGTGAA 345
Qy 138 GtlyvserThrleuValProThrCysIleSerGlnProCysleuProValAlaAnlle 157
Db 344 GGAAAAATCCACTCTAGTCCCTACCTGATTCCTGACCTCTGCTTACCTGTGGCAACATT 285
Qy 158 GtlyProThrArgIleleuProAnleuTyrlleuGtlyCysGlnArgApValleuAnly 177
Db 284 GGCCCTAACCCGAATTTTCCCAATCTTATCTGGCTGCCAGGAGATGCTTCAACAG 225
Qy 178 GtlyleuIleGlnGlnAnnglyIleGtlyTyrlValleuAnAlaserTyThrCysProlys 197
Db 224 GACCTATACAGCAGAAATGGATGGTATGTGTAAATCCAGCTATACCTCTCCCAAG 165
Qy 198 ProAspPheIleProGlnSerHisPheleuArgValProValAnAspSerPheCysGlu 217
Db 164 CCGACCTTATCCCGGCTCTCATTTCTGCGTGCCTGTGAAATGACAGCTTTGTGAG 105
Qy 218 LysIleleuProThrleuAspIlySerValAspPheIleGtlyValAlaserAn 237
Db 104 AAAATTTTCCCGGTGGACAA-TCAGTAGATTTCATTGAGAAAGCAAGCCTC--CAT 48
Qy 238 GtlyCysValleuValHisCysleuAlaglyIleSerArgSerAla 252
Db 47 GGATGTCTTACTGACCTGTTTACCTGGGATCTCCCGCTCCGCC 3

RESULT 14
AK035652 3325 bp mRNA linear HTC 19-SEP-2003
LOCUS AK035652
DEFINITION Mus musculus adult male urinary bladder cDNA, RIKEN full-length
enriched library, clone:9530081F05 product:R10UNCATD MAPK
PHOSPHATASE 7 homolog [Homo sapiens], full insert sequence.
ACCESSION AK035652
VERSION AK035652.1 GI:26330815
KEYWORDS HTC; GAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Carninci,P. and Hayashizaki,Y.
1 High-efficiency full-length cDNA cloning

```

from ORESTERES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

JOURNAL MEDLINE 99279253
REFERENCE 10349686
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL MEDLINE 20499374
PUBMED 11042159
AUTHORS 3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Suni, N., Ishi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishigaki, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL MEDLINE 20530913
PUBMED 11076861
AUTHORS 4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL MEDLINE 11076861
AUTHORS 5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL MEDLINE 11076861
AUTHORS 6
(bases 1 to 3325)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL MEDLINE 11076861
AUTHORS 6
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES SOURCE
Location/Qualifiers
1. .3325
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:9530081F05"
/db_xref="MG1:2399973"
/db_xref="taxon:10090"
/clone="9530081F05"

ORIGIN
Alignment Scores:
Pred. No.: 2,34e-94 Length: 3325
Score: 891.00 Matches: 173
Percent Similarity: 98.87% Conservative: 2
Best Local Similarity: 97.74% Mismatches: 2
Query Match: 56.97% Indels: 0
DB: 11 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x AK035652 (1-3325)
QY 1 METALHISGLWETLIEGLYTHRGHILVVALTHGLUARGLEUVALALEUENGLU 20
DB 460 ATGGCCCATGAGATGATGGAACTCAATTTGTACGAGAGCTGGGCTGCTGGAA 519
QY 21 SGGGLYTHRGHULYSVALLEULEULEASPSERFARPPROPHAYALGIUTYAAANTHR 40
DB 520 AGTGGAGCGAAGAAAGCTGCTGTAATGATGACCGACATTTGGCAATACACTCT 579
QY 41 HISLEUENGLUJALAEANILEASNYSESLYSLEUWETLYSARGLEUGINGLN 60
DB 580 CACATTTTGAAGCATTAATATCACTGCTCCAACTGATGAAGCAAGGTGCAACAG 639
QY 61 ASPLYSVALLEULEUETHTRGULEUULEGINHISERLALYSLSYLSVALAPPIE 80
DB 640 GACAAAGTATTAATTAAGAACTTATCCAACTTGCAGAAAGCAATGACATTGAC 699
QY 81 CYSSEGLINYSVALVALIYTAAPGINSERSEGLINASPYVALIASERLEUSERS 100
DB 700 TGCATATGAGAGGTGGTAGTTATGATCAAGATTCACAGATTTGTTCTGTCTCA 759
QY 101 ASPCYSPHELEUTHRVALLEULEUGLYLYSLEUGLULYSERPHASNSERVALHISLE 120
DB 760 GACTGCTTTCTCACTGTAATTTGGGTAACGAGAGGAGCTTCAACTGTCCACTG 819
QY 121 LEUJALAGLYLYPHEALAGLUPHESERARGCYSPHEROGLYLYUCYSGILUGLY 140
DB 820 CTTCGAGAGTGGCTTGTGTAAGTCTCGTTGTTCCCTGCGCTGTGAGAGAAAGTCC 879
QY 141 THRILEUVALPROTHRYSILSERGLNPROCYSELEUPROVALALASNIIEGLIP 160
DB 880 ACTCTAATCCCTACCTGCAATATCTCACTGCTTGTCTTACCTTTCGAAATGGG 939
QY 161 ARGJLEUPROASNULEUTYRLEUGLYCYGGLNARGASPYVALLEUANSYLS 177
DB 940 CGAATTTTCCCAATCTCTATCTTGGTGCACGAGATGCTCTCAACAG 990
RESULT 15
BU946569/c 792 bp mRNA linear EST 17-OCT-2003
LOCUS 70376910J1 RALUTYXL01 Rattus norvegicus cDNA, mRNA sequence.
DEFINITION BU946569
ACCESSION BU946569
VERSION BU946569.1 GI:37701886
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

AUTHORS

Fu, G.K. and Stuve, L.L.

JOURNAL

An improved method for the construction of cDNA libraries for highly efficient DNA sequencing from the 3' end of expressed genes unpublished (2003)

COMMENT

Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.

FEATURES

source location/Qualifiers

1..792
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/issue_type="Lung"
/clone_lib="RALTUX101"
/note="Rat, Lung, Ciofibrate, 250 mg/kg, 12hr-day 28 sacrifice, SD, M/F, Pool"

ORIGIN

Alignment Scores:

Pred. No.:	4.09e-93	Length:	792
Score:	872.00	Matches:	178
Percent Similarity:	93.78%	Conservative:	3
Best Local Similarity:	92.23%	Mismatches:	9
Query Match:	55.75%	Indels:	4
DB:	13	Gaps:	0

US-10-029-345A-109_COPY_1_302 (1-302) x BU946569 (1-792)

```
Qy      82  SerGlnLySValValValYrAaPGInSerSerGlnAaPValAla-SerLeuSerSerAs 101
      |||
Db      779  AGTCAGAGAGGTGGTAGGTATGATCAGATGCCAGAGATGGTTCGGTCGTCAGACC 720
Qy      101  pCySpHeLeuThValleuLeuGlyLySLeuGluLySerPheAsnSer-ValHisLeuL 121
      |||
Db      719  GGCCTTCCTCAGCTGCTCTCGGTAGCTGAGAGAGAGCTTCAGCTGCTCAGCTGC 660
Qy      121  euAlaGlyLyPheAla-GluPheSerArgCySpHeProGlyLeuCySgIuGlyLySer 140
      |||
Db      659  TTGCAGGTGGGTTCCTGCTGCTCTCGTTCGTTCCCTGCTGCTGAGAGAAATCC 600
Qy      141  ThrLeuValProThrCySleSerGlnProCySleuProValAlaAsnIleGlyProThr 160
      |||
Db      599  ACTCTAGTCCCTACCTGATATCTCAGCCTTGCTTACCTGCTGCTGCTGCTGCTGCT 540
Qy      161  ArgIleLeuProAsnLeuYrLeuGlyCySgInArgAspValleuAsnLySgIuLeuIle 180
      |||
Db      539  CGAATCTCTCCATCTCTATCTCGGCTGCAGAGATGCTCTCAACAAGAGCTGATG 480
Qy      181  GlnGlnAsnGlyIleGlyYrValleuAsnAlaSerYrThrCySpProLyProAspPhe 200
      |||
Db      479  CAACAGAACGGAGTGGCTATGCTAAATGCAAGCAACCTGCTCAAAAGCTGACTTC 420
Qy      201  IleProGluSerHisPheLeuArgValProValAsnAspSerPheCySgIuLySleLeu 220
      |||
Db      419  ATACCGAATCTCAGCTTCCTGCGAGTCCCTGTGATGACAGCTTTTGTGAGAAATCTTG 360
Qy      221  ProTrieuAspLySerValAspPheIleGluYrAlaLySAlaSerAsnGlyCyVal 240
      |||
Db      359  CC-TGGTTGAGACAAGTGTGAGATTTCATTGAGAAACAAGCAAGCTGCTGCTG 301
Qy      241  LeuValHisCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaYrIleMet 260
      |||
Db      300  CTGATTCACGTGCTGGAGATCTCCGCTCGGCAACCAATGCTATGCTATGCTATCATG 241
Qy      261  LysArgMetAspMetSerLeuAspGluAlaYr 271
      |||
Db      240  AAGAGATGACATGCTCTGATGAGGCTTAC 208
```

Search completed: June 21, 2004, 23:33:32
Job time : 2267.97 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 19:56:46 ; Search time 334.168 Seconds
(without alignments)
3839.259 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMIGTQIVTERLVALLE.....PNNFGLGDLDEKRIKQGT 302

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10029345/runcat.21062004.164145.29020/app_query.fasta_1.1294
-DB=genseq.29jand4 -OPMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345.@CGN.1.1.687@runcat.21062004.164145.29020 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1564	100.0	5450	6	ACC60559 Polynucle
2	1564	100.0	5450	6	ACC60572 Polynucle
3	1552	99.2	2118	4	AA630479 Human pro
4	1552	99.2	2732	4	AA609492 Human SGP
5	1552	99.2	3059	6	AA615768 Human sequ
6	1552	99.2	3104	6	ABN59704 Novel hum
7	1552	99.2	3496	6	ABK47596 CDNA enco
8	1552	99.2	3544	5	AA614639 Human CDN

9	1552	99.2	3544	6	ABK49402	ABK49402 CDNA enco
10	1552	99.2	3766	6	ABK14474	ABK14474 Human pro
11	1552	99.2	4790	6	ABN83966	ABN83966 Human gen
12	1552	99.2	5145	5	ABV20833	ABV20833 Human pro
13	1552	99.2	5145	5	ABV21080	ABV21080 Human pro
14	1552	99.2	5145	5	ABV26680	ABV26680 Human pro
15	1552	99.2	5145	5	ABV20978	ABV20978 Human pro
16	1552	99.2	5145	5	ABV21092	ABV21092 Human pro
17	1552	99.2	5145	5	ABV21312	ABV21312 Human pro
18	1552	99.2	5145	5	ABV21316	ABV21316 Human pro
19	1552	99.2	5145	5	ABV26826	ABV26826 Human pro
20	1552	99.2	5145	5	ABV27131	ABV27131 Human pro
21	1552	99.2	5145	5	ABV26923	ABV26923 Human pro
22	1552	99.2	5145	5	ABV27135	ABV27135 Human pro
23	1552	99.2	5145	5	ABV28657	ABV28657 Human pro
24	1552	99.2	5145	5	ABV22827	ABV22827 Human pro
25	1552	99.2	5145	5	ABV26934	ABV26934 Human pro
26	1552	99.2	5145	5	ADA53105	ADA53105 Human cod
27	1544	98.7	2966	4	AAH99685	AAH99685 Human pro
28	1544	98.7	5111	6	ACC60521	ACC60521 Polynucle
29	1482	95.4	2756	6	ACC60560	ACC60560 Polynucle
30	1246	79.7	749	4	AAH06539	AAH06539 Human CDN
31	1194.5	76.4	3332	6	ABK48378	ABK48378 CDNA enco
32	988	63.2	2377	7	ACA64956	ACA64956 Human pro
33	988	63.2	2377	7	ABX10760	ABX10760 Human dua
34	985	63.0	2453	2	AAH86758	AAH86758 CDNA of t
35	985	63.0	2453	2	AAH41236	AAH41236 Murine ne
36	943	60.3	2476	4	AA631013	AA631013 Human dta
37	934.5	59.8	2415	2	AAH86757	AAH86757 CDNA of t
38	742.5	47.5	1171	5	AA872089	AA872089 DNA encod
39	724	46.3	418	5	ABV11252	ABV11252 Human pro
40	724	46.3	461	5	ABV40981	ABV40981 Human pro
41	724	46.3	461	5	ABV22398	ABV22398 Human pro
42	724	46.3	461	5	ABV41327	ABV41327 Human pro
43	715.5	45.7	1917	9	AD807547	AD807547 Novel cod
44	698	44.6	427	5	ABV10887	ABV10887 Human pro
45	635	40.6	419	5	ABV32042	ABV32042 Human pro

ALIGNMENTS

RESULT 1	ACC60559	standard; cDNA; 5450 BP.
ID	ACC60559	
AC	ACC60559;	
XX		
DT	19-JUN-2003 (first entry)	
XX		
DE	Polynucleotide relating to the invention SEQ ID NO: 108.	
XX		
KW	Gene; ss; antiproliferative; hepatocytic; nephrocytic; antitumor;	
KW	antiproliferative; cardiac; cytosolic; gene therapy; liver disease;	
KW	proliferative disorder; renal failure; cardiovascular disorder;	
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;	
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200257460-A2.	
XX		
PD	25-JUL-2002.	
XX		
PF	20-DEC-2001; 2001WO-US050459.	
XX		
PR	20-DEC-2000; 2000US-0256868P.	
PR	30-MAR-2001; 2001US-0280186P.	
PR	01-MAY-2001; 2001US-0287735P.	
PR	05-JUN-2001; 2001US-0295848P.	
XX	25-JUN-2001; 2001US-0300465P.	
XX		
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.	
XX		

PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
PI Krystek S, Mcatee P, Suchard S, Banas D;
XX WPI; 2002-599721/64.
DR P-PSDB; ABR52381.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
PS Claim 1; Fig 13; 801pp; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
SQ
SQ Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.36e-176 Length: 5450
Score: 1564.00 Matches: 302
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-029-345a-109_COPY_1_302 (1-302) x ACC60559 (1-5450)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValIThrGluArgLeuValAlaLeuLeuGlu 20
Db 538 ATGGCCCATGAGTGTGTTGAACTCAATTTTACTGAGAGGTGTGTGCTCTGCGAA 597
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgPropheValGluTyrAsnThrSer 40
Db 598 AGTGAACCGAAAGAGTGTCTGCTAATTGATAGCCGCGCATTTGTGGAATCAATACATCC 657
QY 41 HisIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
Db 658 CACATTTTGGAGCGCATTAATTAATCAACGCTCAAGCTTAATGAAGCGAAAGTTCCAAACAG 717
QY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleValIleValIleValIleVal 80
Db 718 GACCAAGTGTTAATTAACAGAGCTCATCCAGCATTCAGGAACAATAAGTTGACATTGAT 777
QY 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIleValIleVal 100
Db 778 TGCAGTCAAGAGGTGTGTTAGATCAAGATCCCAAGATGTGCTCTCTCTTCA 837
QY 101 AspCysPheLeuThrValIleLeuGluIleValIleValIleValIleValIleValIleVal 120
Db 838 GACTGTTTTCTCACGTGATCTCTGGGTAACTGAGAAAGAGCTTCAACTGTTCACCTG 897
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyIleValIleVal 140
Db 898 CTTCAGAGTGTGTTGCTGAGTGTCTCTGCTGTGTTCCCTGCGCTCTGTGAAGGAATATCC 957
QY 141 ThrLeuValIleProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 958 ACTCTAGTCCCTACCTGATTTCTCAAGCTTGTGCTAATCTGTGTCCCAACATTTGGGCCCAACC 1017
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValIleAsnIleGluLeuIle 180

Db 1018 CGAATTTTCCCAATCTTATCTTGCTGCCAGCGAAGTGTCTCAACAAGAGACTGATA 1077
QY 181 GlnGlnAsnGlyIleGlyTyrValIleAsnAlaSerTyrThrCysProIleProAspPhe 200
Db 1078 CAGCAGAAATGGAGATTTGATGTATGTTAAATCCAGCTATATCCGTCCAAAGCCGACTTT 1137
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIleValIleVal 220
Db 1138 ATCCCCAGTCTCAATTTCTGCTGCTGCTGTGATATGACAGCTTTGTGAGAAATTTTG 1197
QY 221 ProTyrLeuAspIleValIleValIleValIleValIleValIleValIleValIleValIleVal 240
Db 1198 CCGTGTGTGGACCAATCAGTATGATTTTATGAGAAAGCAAAACCTCCATGATGTGTT 1257
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db 1258 CTAGTGCATGTTTAACTGTGAGATCTCCGCTCCGCAACATCCCTATCCCATCATCATG 1317
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluIleValIleValIleVal 280
Db 1318 AAGAGATGGACATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAGAAAGAAAGAAAGAAAG 1377
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuAspTyrGluIleValIleValIleValIleVal 300
Db 1378 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGACTGATGAGAAAGAAAGATTAAGAAC 1437
QY 301 GlnThr 302
Db 1438 CAGACT 1443
RESULT 2
ACC60572
ID ACC60572 standard; cDNA; 5450 BP.
XX
AC 60572;
XX
AC 60572;
XX
DT 19-JUN-2003 (first entry)
XX
DE Polynucleotide relating to the invention SEQ ID NO: 147.
XX
XX Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
XX proliferative disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
OS Homo sapiens.
XX
PN WO200257460-A2.
XX
PD 25-JUL-2002.
XX
PF 20-DEC-2001; 2001WO-US050459.
XX
PR 20-DEC-2000; 2000US-0256868P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0285848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
PI Krystek S, Mcatee P, Suchard S, Banas D;
XX WPI; 2002-599721/64.
DR P-PSDB; ABR52407.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.

tissues expressing PPKR-4) include gastrointestinal (0.385), cardiovascular (0.154), nervous (0.154) and reproductive (0.154). Diseases or conditions associated with tissues expressing PPKR-4 (as a fraction of total tissues expressing PPKR-4) include cancer (0.692), inflammation or trauma (0.308) and cell proliferation (0.231). The encoded protein shows homology to mouse neuronal tyrosine theonine phosphatase 1.1. The invention provides human PPKR-1 to -11 polypeptides (see AAB20322-33) and polynucleotides (see AAF30476-86). It also provides expression vectors, host cells, antibodies, agonists and antagonists, as well as methods for diagnosing, treating or preventing disorders associated with expression of PPKR, including gastrointestinal disorders, immune system disorders, neurological disorders and cell proliferative disorders, including cancer

CC sequences encoding them, as well as various products and methods useful
 CC for the diagnosis and treatment of various phosphatase-related diseases
 CC and conditions. Substance that modulates the activity of phosphatase
 CC polypeptide is used to treat immune-related diseases and disorders,
 CC cardiovascular disease, brain or neuronal-associated diseases and
 CC metabolic disorder, including cancers of tissues, cancers of
 CC hematopoietic origin, diseases of central and peripheral nervous system,
 CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
 CC lateral sclerosis, viral infections, infections caused by prions,
 CC bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual
 CC dysfunction, mood disorders, attention disorders, cognition disorders,
 CC hypertension, psychotic disorder, neurological disorders,
 CC dyskinesias and organ transplant rejection. The present sequence is a DNA
 CC encoding human SGP02 phosphatase polypeptide. This sequence is
 CC classified as dual specificity phosphatase (DSP) and MAP kinase
 CC phosphatase (MKP). SGP02 gene maps to chromosome 12p11.1-p12.1
 CC XX
 SQ Sequence 2732 BP; 710 A; 684 C; 686 G; 652 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,3e-175 Length: 2732
 Score: 1552.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 1
 Query Match: 99.23% Indels: 0
 DB: 4 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AAD09492 (1-2732)

QY 1 MetAlhAGlunecIleGIYThrGlnIleValThrGlnArgLeuValAlaLeuLeuGlu 20
 DB 538 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 537
 QY 21 SerGIYThrGlnValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 DB 598 AGTGAACGGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
 QY 41 HisIleleuGlnIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 60
 DB 658 CACATTTTGGACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 717
 QY 61 AspLysValLeuIleThrGlnLeuIleGlnIleSerAlaIleValIleAspIleAsp 80
 DB 718 GACAAAGTGTATTAACAGAGCTCATCCATTCAGCGAAACATAAGTTGACATTGAT 777
 QY 81 CysSerGlnLysValValValValValValValValValValValValValValVal 100
 DB 778 TGCAGTCAGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
 QY 101 AspCysPheLeuThrValLeuLeuGlyValLeuGlyValLeuGlyValLeuGlyVal 120
 DB 838 GACTGTTTCTCAGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
 QY 121 LeuAlaGlyValPheAlaGlnPheSerArgCysPheProGlyLeuGlyValLeuGly 140
 DB 898 CTTCGAGGTGGTGTGCTGAGTCTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATCC 957
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAlaIleGlyProThr 160
 DB 958 ACTGAGTCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
 QY 161 ArgIleLeuProLeuLeuValLeuGlyValLeuGlyValLeuGlyValLeuGlyVal 180
 DB 1018 CGAATCTCTCCATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCT 1077
 QY 181 GlnGlnAlaGlnIleGlyValValLeuAlaSerValThrCysProLysProAspPhe 200
 DB 1078 CAGCAGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 1137
 QY 201 IleProGlnSerHisPheLeuArgValProValAlaAspSerPheCysGlyValLeu 220
 DB 1138 ATCCCCGAGTCTCATTTCTGCGTGTCTGTGATGACAGCTTTTGTGAGAAATTTTG 1197

QY 221 ProTrpLeuAspLysSerValAspPheIleGlyValAlaLysAlaSerAsnGlyCysVal 240
 DB 1198 CCTGGTTTGACAAATACATGATTTTCATTTGAGAAAGCAAAAGCCCTCAATGATGTGT 1257
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
 DB 1258 CTAGTCACATGTTTACCTGGATCTCCCTCCCTCCGACACATCGCTATCGCTCAATCATG 1317
 QY 261 LysArgMetAspMetSerLeuAspGlnAlaValArgPheValIleGlyValArgProThr 280
 DB 1318 AAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 1377
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysValIleLysAsn 300
 DB 1378 ATRTCTCCAACTCAATTTTCTGGCCCAACTCTGACTATAGAGAAAGATTAAAGAAC 1437
 QY 301 GlnThr 302
 DB 1438 CAGACT 1443

RESULT 5
 AAS15768
 ID AAS15768 standard; cDNA; 3059 BP.
 XX
 AC AAS15768;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE cDNA sequence of human (dual specificity phosphatase) DUSP-10.
 XX
 KW Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;
 KW neuronal degeneration syndrome; Alzheimer's disease; depression;
 KW schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
 KW osteoporosis; diabetes; gene therapy; chromosome 12; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 127..2124
 FT /tag= a
 FT /product= "DUSP-10 protein"
 XX
 PN MO200177340-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-EP003966.
 XX
 PR 10-APR-2000; 2000EP-00107143.
 XX
 PA (MERE) MERCK PATENT GMBH.
 XX
 PI Duecker K;
 XX
 DR WPI; 2002-010917/01.
 DR P-PSDB; AAD09946.
 XX
 PT Novel dual specificity phosphatase polypeptides useful for treating
 PT cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
 PT disease, depression, schizophrenia, asthma and immune disorders.
 XX
 PS Claim 5; Page 34-37; 43pp; English.
 XX
 CC The present invention relates to a new isolated dual specificity
 CC phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
 CC sequence that is fully defined in the specification. The invention also
 CC provides a sequence encoded by a 3059 nucleotide sequence fully defined
 CC in the specification, and a sequence having at least 95 % identity to the
 CC polypeptide, or fragments or variants of DUSP-10. The invention is useful
 CC for treating cancer e.g. leukaemia, colon carcinoma, lung cancer,
 CC prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
 CC stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
 CC schizophrenia, cardiac myotrophies, asthma, immune disorders,

CC inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
CC osteoporosis, diabetes and diabetes associated diseases. The molecules of
CC the invention are also useful as vaccines for inducing immunological
CC response in a mammal, in disease diagnosis and in assays for screening
CC agonistic or antagonistic compounds. Other uses of the invention include
CC identifying membrane bound or soluble receptors, as a diagnostic reagent,
CC in chromosome localisation studies, and as a valuable tool in tissue
CC expression studies. The present sequence represents cDNA of the human
CC dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome
XX 12

XX Sequence 3059 BP; 831 A; 761 C; 709 G; 758 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.54e-175	Length:	3059
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345a-109_COPY_1_302 (1-302) x AAS15768 (1-3059)

```
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 127 ATGGCCCATGAGATGATGAGTGAATGCAATGTTTACTGAGAGGTTGGCTGCTGCGGAA 186
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 187 AGTGAACGGAAGGAGGCTGCTAATGATGAGGCGGCTTTGTGATATCAATCAATTC 246
QY 41 HisIleLeuGluAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 60
Db 247 CACATTTTGGAGCCCATTAATATCACTCTCCAGCTTATGAACGAAAGTTGCAACAG 306
QY 61 AspIleValLeuIleThrGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Db 307 GACAAAGTGTATTAACAGAGCTCATCCGATTCAGGCAACATTAAGGTTGACATTTGAT 366
QY 81 CysSerGluValValValValValValValValValValValValValValValValVal 100
Db 367 TCGAGTCAGAGAGTGTGATGATTCATCAAGCTCCCAAGATGTTGCTCTCTCTCTCA 426
QY 101 AspCysPheLeuThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
Db 427 GACTGTTTCTCACTGATCTTCTGGGTTAACTGGAGAAAGCTTCACCTGTTCCACTG 486
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLeuSer 140
Db 487 CTTCGAGGTGGGTGGTGGAGTTCCTGCTGTTTCCCTGCTCTGTGAAGGAAATTC 546
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAlaIleGlyProThr 160
Db 547 ACTCTAGTCCCTACTGCAATTTCTCAGCTTGTACTCTGTGCAACATTTGGGCAAC 606
QY 161 ArgIleLeuProAsnLeuTyLeuGlyCysGlnArgAspValLeuAsnLeuGluLeuLeu 180
Db 607 CGAATTTCTCCCAATCTTATCTTGGCTGCGCAAGCAATGCTTCAACAAAGAGCTGATG 666
QY 181 GlnGlnAsnGlyIleGlyTyValLeuAsnAlaSerTyThrCysProLysProAspPhe 200
Db 667 CAGCAGAAATGGAGTGTATGTTAAATCCGCAATACCTGTCCAAAGCCTGACTTT 726
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlyValLeu 220
Db 727 ATCCCCGAGTCTCATTTCTGCGTGTGCTGTGATATGACAGCTTTGTGGAAGAAATTTG 786
QY 221 ProTrpLeuAspLysSerValAspPheIleGluValAlaLysAlaSerAsnGlyCysVal 240
Db 787 CCGGTGTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGTT 846
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
```

```
Db 847 CTATGCACTGTTAGTGGATCTCCGCTCCGCCACCATGCTATGCTACATCATG 906
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyTrpPheValLysGlyValArgProThr 280
Db 907 AAGAGATGGACATGCTCTTTGATGAAAGCTTACAGATTTGGAAAAAGAAAGCTTACT 966
QY 281 IleSerProAsnPheAsnPheLeuGlyGluLeuLeuAspTyTrpGlyLysIleLysAsn 300
Db 967 ATATCTCCAACTCAATTTTCTGGCCCACTCTGAGCTATGAGAAAGATTAAAGAAC 1026
QY 301 GlnThr 302
Db 1027 CAGACT 1032
```

RESULT 6
ABN59704
ID ABN59704 standard; cDNA; 3104 BP.
XX
XX
AC ABN59704;
XX
XX
DT 28-JUN-2002 (first entry)
XX
XX

Novel human coding sequence SEQ ID NO: 115.

Human; anti-neurotic; vulnerary; anti-inflammatory; immunomodulator;
KW antifertility; cerebroprotective; cytosolic; rheumatic; gene therapy;
KW neuroprotective; anti-parinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

XX 10-SEP-2001; 2001WO-US026015.

XX 11-SEP-2000; 2000US-00659671.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI; 2002-292408/33.

XX P-P-SDB; ABB97291.

PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.

PS Claim 1; SEQ ID NO 115; 509bp; English.

XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to help tissue
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat
XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a coding sequence of the
XX invention

XX Sequence 3104 BP; 845 A; 772 C; 718 G; 769 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.58e-175	Length:	3104
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345a-109_COPY_1_302 (1-302) x ABN59704 (1-3104)

```

QY      1 MetAlAHiegluMetilegIyThrgInileValThrgIuArgleuValAlaleuLeuglu 20
Db      159 ATGGCCCATAGATGATGGAACTCAATTTGTTACTGAGAGGTTGGCTGCTGCTGGAA 218
QY      21 SerGIyThrgIuArgleuValleuLeuileAspSerArgProhpeValGluYrAnThSer 40
Db      219 AGTGGAAAGGAAAGAGTCTGCTATGATGATGCGGCGCATTTGGGAATCAATACATCC 278
QY      41 HisIleuGlulAlaleuAlaleuAlaenIleAsnCySerIySerMetIySArgArgleuGln 60
Db      279 CACATTTTGAAGCATTAATATCACTGCTCCAGAGTTTGAAGAGAGTTTGAAGAGAG 338
QY      61 AsplYValleuileThrgIuLeuileGlnHisSerAlalyshIstlyValAsp1leAsp 80
Db      339 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCACGAAACATTAAGTTGACATTGAT 398
QY      81 CySerGlnIyValValValValTyAspGlnSerSerGlnAspValAlaSerIleuSerSer 100
Db      399 TGAAGTCAGAGGTTGATGTTACGATCAAGAGCTCCAGAGATGTTGCTCTCTCTTCA 458
QY      101 AspCySerPheleuThrValleuLeuGlyLyAspGluIySerSerPheAsnSerValHisleu 120
Db      459 GACGTGTTTCTCACTGACTTCTGGGTAACCTGGAGAGAGAGCTTCACATCTGTTCACTG 518
QY      121 LeuAlaGlyGlyPheAlaGluPheSerArgCySerPheProGlyIyLeuCyGluGlyIySer 140
Db      519 CTGCGAGTGGGTTGCTGAGTCTCTGTTGTTCCCTGCTGCTGAGAGAGAAATCC 578
QY      141 ThrleuValProthrCySileSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
Db      579 ACTGTAGTCCCTACCTGATTTTCACAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638
QY      161 ArgIleuProAsnleuIyLeuGlyCyGlnArgAspValleuAsnIyGluLeuIle 180
Db      639 CGAATCTTCCCAATCTTTACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698
QY      181 GlnGlnAsnGlyIleGlyIyValleuAsnIleSerTyThrCyProIyProAspPhe 200
Db      699 CACGAGATGGATGGTATGTTGTTAAATGCCAGCAATCTGCTCAAGAGCTGACTTT 758
QY      201 IleProGlnSerIyPheleuArgValProValAsnAspSerPheCyGluIyIleu 220
Db      759 ATCCCGAGTCTCAATTTCTGCGGTCGCTGATGATGACGCTTTGTGAGAAATTTTG 818
QY      221 ProTripleuAspIySerSerValAspPheIleGluIyValIyAsnIleSerAsnGlyCyVal 240
Db      819 CCGTGTGGACAAATCAATGATTTCAATTCAGAAAGCAAAAGCTTCATGATGATGTT 878
QY      241 LeuValHisCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIyIleMet 260
Db      879 CTAGTGACATGTTAGCTGGATCTCCGCTCGCCACATCCCTATCGCTCATCATCATG 938
QY      261 LysArgMetAspMetSerIleuAspGluAlaTyArgPheValIyGluIySArgProThr 280
Db      939 AAGAGATGAGCAATGCTTTAGATGAGAGCTTACAGATTTGTGAAGAAAAAAGACCTACT 998
QY      281 IleSerProAsnPheAsnPheleuGlyGlnleuLeuAspTyArgIySlyIleIyAsn 300
Db      999 ATATCTCCCAACTTCATTTTCTGGGCAACTCTGAGACTATGAGAGAGATTAAGAAC 1058
QY      301 GlnThr 302
Db      1059 CAGACT 1064

```

RESULT 7

ABK47596
ID ABK47596 standard, cDNA; 3496 BP.
AC ABK47596;
XX

```

DT      02-JUL-2002 (first entry)
XX      DE      cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.
XX      KW      Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
XX      KW      mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
XX      KW      cancer; graft-versus-host disease; allergy; metabolic disease;
XX      KW      abnormal cell growth; abnormal cell proliferation; contact inhibition;
XX      KW      cell cycle abnormality; anchorage independent cell growth; apoptosis;
XX      KW      intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; 89.
XX      OS      Homo sapiens.
XX      FH      Key      Location/Qualifiers
XX      FT      CDS      562..2559
XX      FT      /tag= a
XX      FT      /product= "Human dual-specificity phosphatase-3 (DSP-16)
XX      FT      protein"
XX      PN      WO200226997-A2.
XX      PD      04-APR-2002.
XX      PF      25-SEP-2001; 2001WO-US030124.
XX      PR      26-SEP-2000; 2000US-0235487P.
XX      PA      (CEPT-) CEPTYR INC.
XX      PI      Luche RM, Wei B;
XX      DR      WPI: 2002-315802/35.
XX      DX      P-PSDB; AAU79156.
XX      PT      New DSP-16 polypeptide, useful for identifying modulators of its
XX      PT      activity, which can be used in the treatment of disorders such as
XX      PT      Duchenne muscular dystrophy, or cancer.
XX      PS      Claim 7; Fig 1; 87bp; English.
XX      CC      The present invention relates to a new polypeptide, DSP-16, having a 665
XX      CC      amino acid sequence, given in the specification, or a variant having at
XX      CC      least 50 % identical residues, which retains the ability to
XX      CC      dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
XX      CC      invention can be used for identifying agents which modulate DSP-16
XX      CC      activity, for modulation of a proliferative response in a cell, survival
XX      CC      of a cell, or differentiation of a cell. The cell displays contact
XX      CC      inhibition of cell growth or anchorage independent growth and may display
XX      CC      altered intercellular adhesion. The agent may modulate apoptosis, or the
XX      CC      cell cycle. The identified modulators can be used to treat Duchenne
XX      CC      muscular dystrophy, cancer, graft-versus-host disease, autoimmune
XX      CC      diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
XX      CC      cell proliferation, and cell cycle abnormalities. The present nucleic
XX      CC      acid sequence is that of the human dual-specificity phosphatase-3 (DSP-
XX      CC      16) gene located on chromosome 12p. This sequence encodes the human DSP-
XX      CC      16 protein of the invention
XX      SQ      Sequence 3496 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;

```

Alignment Scores: 1 9e-175 Length: 3496
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 6 Gaps: 0

US-10-029-345a-109_COPY_1_302 (1-302) x ABK47596 (1-3496)

```

QY      1 MetAlAHiegluMetilegIyThrgInileValThrgIuArgleuValAlaleuLeuglu 20
Db      562 ATGGCCCATAGATGATGGAACTCAATTTGTTACTGAGAGGTTGGCTGCTGCTGGAA 621

```

QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 DB 622 AGTGAACGGAAGAAAGTCTCTTAATGATGATGCGGCAATTTGGTAATCAATATCAATCC 661
 QY 41 HisIleLeuGluAlaIleAsnIleAsnCySerSerLysLeuMetLysArgArgLeuGlnGln 60
 DB 682 CACATTTTGGAAAGCAATTAATATCAATCACTGCTCCAACTTATGAAGCAAGCTTGCACACG 741
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerLysAlaLysHisLysValAspIleAsp 80
 DB 742 GACAAAGCTTAATTAACAGACTCATCCAGCAATTCAGCAAAACATTAAGTGAATGATGAT 801
 QY 81 CysSerGlnLysValValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 802 TGCAGTCAGAAAGGTTGATGATTAAGATCAAAAGCTCCCAAGATGCTCTCTCTCTTA 861
 QY 101 AspCysPheLeuThrValLeuLeuGluLysLeuGluLysSerPheAsnSerValHisLeu 120
 DB 862 GACTGTTTCTCACTGACTTCTGGGTAACTGGAAGAGCTTCAACTCTGTTCACTG 921
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGluGlyLysSer 140
 DB 922 CTTCAGAGTGGGTTTGCTGAGTTCTCTCGTTGTTTCCCTGCTCTGTGAAGAAATCC 981
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 982 ACTTAGTCCCTTAACGCAATTTCTCAAGCTTGTCTTACCTGCTTGCACCACTGGGCAACC 1041
 QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
 DB 1042 CGAATTTCTCCCAATCTTATATCTGCTGCTCCAGCAAGATGCTCTCAACAAAGAGCTGATG 1101
 QY 181 GlnGlnAsnGlyTyrIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
 DB 1102 CAGCAAAATGGGATGTTATGTTAAATGCACAGAAATACCTGTCCAAAGCTGACTTT 1161
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
 DB 1162 ATCCCGAATCTCATTTCTCGGTGCTGCTGGAATGACAGCTTTTGTGAAGAAATTTTG 1221
 QY 221 ProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
 DB 1222 CCGTGTGTCAGAAATCACTAGATTTCAATTGGAAGCAAAAGCTCCAAAGATGATGTT 1281
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleTyrIleMet 260
 DB 1282 CTAGTGCACTGTTAGCTGGATCTCCGCTCCGCCAACATCGCTATCGCTACATCATG 1341
 QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
 DB 1342 AAGAGGATGACATGCTTTAGTGAAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 1401
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysValIleLysAsn 300
 DB 1402 AATCTCCAAATCTTAATTTCTGGGCCAATCTCTGCACTATGAAGAAAGATTAAGAAC 1461
 QY 301 GlnThr 302
 DB 1462 CAGACT 1467
 RESULT 8
 AAS14639
 ID AAS14639 standard; cDNA; 3544 BP.
 AC AAS14639;
 AC AAS14639;
 DT 18-DEC-2001 (first entry)
 XX
 DE Human cDNA encoding dual specificity phosphatase 21117.
 XX
 KW Human; ss; dual specificity phosphatase 21117; hepatotropic; cytostatic;
 KW hematopoietic disorder; autoimmune disorder; diabetes mellitus;
 KW rheumatoid arthritis; multiple sclerosis; Crohn's disease;

KW liver disorder; erythroid associated disorder; hemolytic anaemia;
 KW cellular proliferative; differentiative disorder; leukaemia;
 KW acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 589..2586
 FT /tag= a
 FT /product= "Dual specificity phosphatase 21117"
 PN MO20013059-A2.
 PD 04-OCT-2001.
 PF 23-MAR-2001; 2001WO-US009477.
 PR 24-MAR-2000; 2000US-0191858P.
 PA (MILL-) MILLENNIUM PHARM INC.
 PI Meyers RA;
 XX WPI; 2001-611635/70.
 DR P-PSDB; AAU09016.
 XX
 PT New human dual specificity polypeptides and nucleic acids for diagnosis
 PT of disease and treatment of e.g. liver disorders.
 PS Claim 1; Fig 1; 143pp; English.
 XX
 CC The invention relates to two novel human dual specificity phosphatases
 CC designated 21117 and 38692, the nucleic acids encoding them (including
 CC fragments, allelic variants, their complements or nucleic acids that
 CC hybridise to them) and antibodies raised against the proteins. The
 CC antibody is useful for detecting the presence of the polypeptide, and the
 CC nucleic acid fragments are useful for detecting the presence of the
 CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
 CC antisense sequences) are useful for modulating the activity or expression
 CC of the polypeptide/nucleic acid, useful for the treatment of e.g.
 CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
 CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
 CC listed in the specification) liver disorders, erythroid associated
 CC disorders (e.g. haemolytic anaemia) cellular proliferative or
 CC differentiative disorders, leukaemias (e.g. acute myeloid leukaemia),
 CC carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
 CC 38692 are also useful for modulating the proliferation, survival,
 CC migration or differentiation of a 38692 or 21117-expressing cell. The
 CC polypeptide and nucleic acids are useful for identifying modulating
 CC agents. The present sequence encodes the dual specificity phosphatase
 CC 21117
 XX
 SQ Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 1,94e-175 Length: 3544
 Score: 1552.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 1
 Query Match: 99.23% Indels: 0
 DB: Gaps: 0
 US-10-029-345A-109_COPY_1_302 (1-302) x AAS14639 (1-3544)
 QY 1 MetaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
 DB 589 ATGGCCATGAGATGATTTGGAATCAATTTCTACTAGAGCTTGGCTCTGCTGGAA 648
 QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 DB 649 AGTGAACGGAAGAAAGTCTCTTAATGATGATGCGGCAATTTGGTAATCAATATCAATCC 708
 QY 41 HisIleLeuGluAlaIleAsnIleAsnCySerSerLysLeuMetLysArgArgLeuGlnGln 60

```
Db 709 CACATTGGAGCCATTATATCACTGCTCCAGCTTATGAAAGCAAGCTTGCAACAG 768
Qy 61 AepLyVaLeu1leThrg1Leu1leGlnH1SerAla1ySh1b1yVa1Aap11eap 80
Db 769 GACAAAGTGTATTAATACAGAGCTATCCAGCATTCAGCAAAATAGTGAAGTGAATGAT 828
Qy 81 CysSerGln1yVa1Va1Va1TyraepGlnSerSerGlnAap11aSer1eapSer 100
Db 829 TGCAGTCAGAAAGTGTATTAATACAGATCAAGCTCCCAAGATGTGCTCTCTCTCA 888
Qy 101 AspCysPheLeuThrVal1LeuLeuGly1yVa1SerPheAenSerVal1H1eLeu 120
Db 889 GACTGTTTCTCACTGACTTCTGGGTAACCTGGAAGAGCTTCAACTCTGTTCACTG 948
Qy 121 LeuAlaGly1yVa1AaGlnPheSerArgCysPheProGly1eucYsg1yLy1ySer 140
Db 949 CTTCAGAGTGGTGTGCTGAGTCTCTGTTGTTCCCTGCTGCTGTAAGAAATCC 1008
Qy 141 ThrLeuVal1ProThrCys1leSerGlnProCysLeuProVal1Aaen1leGlyProThr 160
Db 1009 ACTCTAGTCCCTTACCTGCACTTCTCAGCCTTGTCTTACCTGTGCAACATTTGAGCCAAACC 1068
Qy 161 Arg1leuProAenLeu1yLy1eGlnArgAapVal1eAen1yLy1eLeu1le 180
Db 1069 CGAATCTTCCCAATCTTATCTTGCTGCTGCAAGATGCTCTCAACAAAGAGCTGATG 1128
Qy 181 GlnGlnAaenGly1leGly1yVa1LeuAaen1leSer1yThrCysPro1yProAapPhe 200
Db 1129 CACAGAAATGGATGGTGTATGTTAAATGACAGCAATCTGCTCAAGCTGACTTT 1188
Qy 201 11aProGlnSer1yPheLeuArgVal1ProVal1AaenAapSerPheCysGly1yLy1eLeu 220
Db 1189 ATCCCGAGTCTCAATCTCTGCGTGTGCTGTAATGACAGCTTTGTGAAAAATTTTG 1248
Qy 221 Pro1yThrLeuAap1ySerVal1AapPhe1leGly1yVa1Aaen1yLy1eLeu 240
Db 1249 CCTGTGTGCAAAATCAATGATTTCTTATGAGAAAGCAAAAGCTTCAATGATGTGTT 1308
Qy 241 LeuVal1H1eCysLeuAaGly1leSerArgSerAlaThr1leAa1leAa1y1leMet 260
Db 1309 CTAGTGCACATGTTAGTGGATCTCCGCTCGGCAACATGCTATGCTCATCATCATG 1368
Qy 261 LyArgMetAapMetSerLeuAapGly1Aa1yThrArgPheVal1yGly1yVa1yAapProThr 280
Db 1369 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAGACCTACT 1428
Qy 281 11leSerProAenPheAenPheLeuGly1GlnLeuLeuAap1yThrGln1yLy1yLeuAaen 300
Db 1429 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGAGCTATGAGAAAGATTAAGAAC 1488
Qy 301 GlnThr 302
Db 1489 CAGACT 1494
RESULT 9
ABK49402
ID ABK49402 standard; cDNA; 3544 BP.
XX
XX ABK49402;
XX
XX
XX 02-JUL-2002 (first entry)
XX
XX cDNA encoding human dual specificity phosphatase 21117 protein.
XX
XX Human: dual specificity phosphatase 21117; erythroid-related disorder;
XX haemotopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
XX erythrocytosis; liver-related disorder; cancer; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX
XX FT CDS 589..2586
```

```
PT /**tag= a
FT /product= "Human dual specificity phosphatase 21117"
PT /note= "Specifically claimed in claim 2"
XX
XX US2002034807-A1.
XX
XX 21-MAR-2002.
XX
XX 23-MAR-2001; 2001US-00816494.
XX
XX 24-MAR-2000; 2000US-0191858P.
XX
XX (MEYE/) MEYERS R A.
XX
XX Meyers RA;
XX
XX MPI; 2002-351088/38.
XX
XX P-PSDB; AAU79929.
XX
XX
XX PT New nucleic acids, designated 38692 and 21117, encoding dual specificity
XX phosphatases for treating cell proliferation and differentiation
XX disorders including hematopoietic and erythroid-related disorders and
XX cancers.
XX
XX Claim 2; Fig 1; 76pp; English.
XX
XX
XX The present invention relates to new nucleic acids designated 38692 and
XX 21117 encoding dual specificity phosphatase family members. The nucleic
XX acid, polypeptide encoded by it, and antibody specific for the
XX polypeptide may be used to diagnose and treat hematopoietic-related
XX disorders such as leukemias and autoimmune diseases, erythroid-related
XX disorders such as anemias and erythrocytosis, liver-related disorders,
XX and cancers, particularly of the breast, colon, adipose, prostate and
XX lung. The present nucleic acid sequence encodes the human dual
XX specificity phosphatase 21117 protein of the invention, as described
XX above
XX
XX SQ Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1,94e-175 Length: 3544
XX Score: 1552.00 Matches: 300
XX Percent Similarity: 99.67% Conservative: 1
XX Best Local Similarity: 99.34% Mismatches: 1
XX Query Match: 99.23% Indels: 0
XX DB: Gaps: 0
XX
XX US-10-029-345A-109_COPY_1_302 (1-302) x ABK49402 (1-3544)
XX
XX Qy 1 Meta1aH1eGlnMet1leGly1Thrg1n1leVal1Thrg1uArgLeuVal1AaenLeuGln 20
XX 589 ATGGCCCATGAGATGATGGAATCAAAATGTTACTGAGAGGTGGTGGCTGTGGAA 648
XX
XX Qy 21 SerGlyThrGln1yVa1LeuLeu1eAapSerArgProPheVal1Gln1yThraenThrSer 40
XX 649 AGTGAACGAAAGAAAGTGTGCTGCTTATGATGAGCGGCAATTTGTGAAATCAATACATCC 708
XX
XX Qy 41 H1leLeuGln1Aa1leAaen1leAenCysSer1yLeuMetLyAArg1yGlnGln 60
XX 709 CACATTTTGGAGCCATTATATCAACTGCTCCAGACTTATGAGAGAAAGTTGCAACAG 768
XX
XX Qy 61 AepLyVa1Leu1leThrg1Leu1leGlnH1SerAla1ySh1b1yVa1Aap11eap 80
XX 769 GACAAAGTGTATTAATACAGAGCTATCCAGCATTCAGCAAAATAGTGAAGTGAATGAT 828
XX
XX Qy 81 CysSerGln1yVa1Va1Va1TyraepGlnSerSerGlnAap11aSer1eapSer 100
XX 829 TGCAGTCAGAAAGTGTATTAATACAGATCAAGCTCCCAAGATGTGCTCTCTCTCA 888
XX
XX Qy 101 AspCysPheLeuThrVal1LeuLeuGly1yVa1SerPheAenSerVal1H1eLeu 120
XX 889 GACTGTTTCTCACTGACTTCTGGGTAACCTGGAAGAGCTTCAACTCTGTTCACTG 948
```

```
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGluLeuGlyGlySer 140
Db 949 CTGGAGAGTGGGTTGGTGAAGTCTCTCGTGTTTCCCTGGCTCTGTGAAGAAATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAenIleGlyProThr 160
Db 1009 ACTTAGTCCTTACTGATTTCCAGGCTTGCTTACTGTTGGCAACATTGGGCAACC 1068
QY 161 ArgIleLeuProAenLeuTyrLeuGlyCysGlnArgSpValLeuAenIleGluLeu 180
Db 1069 CGAATTCCTCCCAATCTTATCTTGCTGGCCAGAGATGTCCTCAACAAAGACTATG 1128
QY 181 GlnGlnAenGlyIleGlyTyrValLeuAenAlaSerTyrThrCysProIysProAspPhe 200
Db 1129 CAGAGAAATGGAGATTGTTATGTTAAATCCAGCAATACCTGTCCAAAGCTGACTTT 1188
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
Db 1189 ATCCCGAGTCTCATTCTCTGCTGCTGATGATGACAGCTTTGTGAGAAATTTTG 1248
QY 221 ProTyrLeuAspIysSerValAspPheIleGluIlyAlaIlyAlaSerAenGlyCysVal 240
Db 1249 CCGTGCTTGGACAATCAGTGAATTCATTGAGAAAGCAAGCCCTCCAAATGGATGTT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerIleThrIleAlaIleAlaTyrIleMet 260
Db 1309 CTAGGCACTGTTTACTGCTGGATCTCCGCTCCGACCACTGCTATGCTTACTCATG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIlyGluIlyArgProThr 280
Db 1369 AAGAGATGAGCATGCTTTAGATGACCTTACAGATTGTGAAAGAAAGAAAGACTTACT 1428
QY 281 IleSerProAenPheAsnPheLeuGlyGlnLeuLeuAspTyrGluIlyValIleIlyAsn 300
Db 1429 ATATCTCCAAACTTCAATTTCTGGGCCAAGCTCTGAGCTATGAGAGAAAGATTAGAAC 1488
QY 301 GlnThr 302
Db 1489 CAGACT 1494
RESULT 10
ABK14474 ID ABK14474 standard; cDNA; 3766 BP.
AC ABK14474;
DT 08-MAY-2002 (first entry)
DE Human protein phosphatase 7 (PP7) cDNA sequence.
XX Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
XX acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
XX Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
XX dementia; Parkinson's disease; developmental disorder; Down's syndrome;
XX cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
XX melanoma; myeloma sarcoma; gene; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 538..2535
FT /tag= a
FT /product= "Protein_phosphatase_7 (PP7) "
```

```
PR 18-AUG-2000; 2000US-0226728P.
PR 30-AUG-2000; 2000US-0229254P.
PR 08-SEP-2000; 2000US-0231366P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE,
PI Stewart BA, Gandhi AR, Patterson C, Lee EA, Hatalla AJA, Lu DM,
PI Tribouley CM, Griffen JB, Baughn MR, Yue H, Warren BA, Nguyen DB,
PI Malia NK, Kearney L,
XX WPI; 2002-188735/24.
DR P-PSDB; AAU75789.
XX
XX New protein phosphatases, useful for diagnosing, treating or preventing
PT immune system disorders (e.g. Crohn's disease), neurological disorders
PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
PT cancers).
XX
PS Claim 5; Page 114-115, 117pp; English.
```

The present invention relates to a new polypeptide, a naturally occurring amino acid sequence at least 95 % identical to it, a biologically active fragment of it or an immunogenic fragment of it. The polypeptides, polynucleotides, agonists and antagonists are useful for diagnosing, treating or preventing disorders associated with aberrant expression of protein phosphatases (PP), particularly immune system disorders e.g. acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia, asthma or Crohn's disease, neurological disorders e.g. epilepsy, Huntington's disease, dementia or Parkinson's disease, developmental disorders e.g. Down's syndrome, or cell proliferative disorders e.g. cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma or sarcoma. The present nucleic acid sequence encodes human protein phosphatase 7 (PP7) which is one of several human protein phosphatases (AAU75783-AAU75792) of the invention

Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,13e-175	Length:	3766
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345A-109_COPY_1_302 (1-302) x ABK14474 (1-3766)

```
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
Db 538 ATGGCCCATGAGATGGATTTGGAACCTCAATTTGTTACTGAGAGTTGGCTGCTGGAA 597
QY 21 SerGlyThrGluIysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db 598 AGTGAACGGAAAAGTCTGCTTAATGATAGCCGGCCATTTGGATGAATCAATATCC 657
QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIysLeuMetIysAlaGluLeuGln 60
Db 658 CACATTTTGGAAAGCCATTAAATATCACTGCTCCAAAGCTTATGAAAGCAAGGTTGCCAACG 717
QY 61 AspIysValLeuIleThrGluLeuIleGlnHisSerIlyAlaIlyHisValAspIleAsp 80
Db 718 GACAAAGTTTAAATTACAGAGCTATCCAGCATTCAGCGAAACATAGAGTTGACATTGAT 777
QY 81 CysSerGluIysValIleValIlyTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 778 TGCAGTCAGAAAGTTTGAAGTTTACGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 837
QY 101 AspCysPheLeuThrValLeuLeuGlyIlyIysLeuGluIlySerPheAsnSerValHisLeu 120
Db 838 GACTGTTTTCACCTGATCTCTGGTTAACTGAGAAAGAGCTTCAACTCTGTTCCACTG 897
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGluLeuGlyGlySer 140
```



```
DB 898 CTTGCGAGTGGGTTTCTAGTCTCTCGTGTTCCTCGGCTCTGTGAAGAAAATCC 957
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaIleuIleGlyProThr 160
DB 958 ACTGTGTCCTCACTGACATTTCTCGAGCCCTGCTTACCTGTTCCAAACATTTGGCCAAACC 1017
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleuGlyLeu 180
DB 1018 CGAATCTTCCCAATCTTATCTTGCTGCGACGAGATGCTCTCAACAGAGCTGATG 1077
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProIleProAspPhe 200
DB 1078 CAGCAGAAATGGGATTGGTTATGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 1137
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
DB 1138 ATCCCGAGCTCATTTCTCGCTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG 1197
QY 221 ProTrpLeuAspLysSerValAspPheIleGluIleValAlaIleSerAsnGlyCysVal 240
DB 1198 CCGTGGTGGACAAATCAGTAGATTTTCATTGAGAAAAGCAAAAGCTCCAAATGGATGTT 1257
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
DB 1258 CTAGTGCACCTGTTAGCTGGGATCTCCGCTCCGCCACCATGCTATCGCTTACATCATG 1317
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluIleValArgProThr 280
DB 1318 AAGAGAGATGACATGCTTATGATGAAGCTTACAGATTGTGTAAGAAAAGAAAGACCTACT 1377
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluIleValIleIleAsn 300
DB 1378 AATCTCCAAACTTCAATTTCTGGGCCAACTCTGAGCATATGAGAAAGATTAAAGAAC 1437
QY 301 GlnThr 302
DB 1438 CAGACT 1443

RESULT 11
ABN83966
ID ABN83966 standard; DNA; 4790 BP.
XX
AC ABN83966;
XX
DT 06-SEP-2002 (first entry)
XX
DE Human gene sequence #13.
XX
KW Human; brain; tonsil; hippocampus; foetal brain; diagnosis; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 184..2181
FT /tag= a
XX
XX WO200252005-A1.
XX
XX 04-JUL-2002.
XX
XX 20-DEC-2001, 2001MO-JP011217.
XX
XX 22-DEC-2000, 2000JP-00389742.
XX
XX (KAZU-) KAZUSA DNA RES INST FOUND.
XX (CELE-) CELESTAR LEXICO-SCI LTD.
XX
XX Ohara O, Nagase T, Nakajima D;
XX
XX MPI; 2002-500762/53.
XX
XX P-PsDB; ABB97946.
XX
```

```
PT Genes and their expression products cloned from human cDNA libraries for
PT treatment and diagnosis of diseases associated with their expression.
XX
XX Claim 1(a); Page 111-117; 238pp; Japanese.
XX
XX The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies,
CC oligonucleotide probes and DNA chips for diagnosis and identification of
CC drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABN83954-ABN83984 represent
CC human gene sequences of the invention
XX
XX Sequence 4790 BP; 1293 A; 1064 C; 1152 G; 1281 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3, 08e-175 Length: 4790
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 6 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x ABN83966 (1-4790)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 184 ATGGCCCATGAGATGATGGAAGTCAATCAATTTGTTACTGAGAGGTGTGCTGTGAA 243
QY 21 SerGlyThrGluIleValLeuLeuIleAspSerArgProPheValGluIleAsnThrSer 40
DB 244 AGTGAACCGAAAAAAGTGTGCTTAATGTATGATGCGGCCCATTTGTGTAATCAATATCATCC 303
QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerTyrLeuMetCysArgArgLeuGln 60
DB 304 CACATTTGGAAAGCCATTAAATCAACTCTCCAGCTTATGAGAGGATTGCAACAG 363
QY 61 AspLysValIleuIleThrGluLeuIleGlnHisSerAlaLysHisIleValAsp 80
DB 364 GACAAAGTGTAAATTAACAAGCTCATCAGCATTCAGCAATTCAGCAATAGGTGACATTGAT 423
QY 81 CysSerGlnLysValValIleValIleAspGlnSerSerGlnAspValAlaSerLeuSer 100
DB 424 TGCAGTCAGAAAGTGTGATGTTACGATCAAAAGCTCCCAAGATGTTCCCTCTCTTCA 483
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
DB 484 GACTGTTTCTCACTGACTTCTCGGTTAACTGAGAGAGCTTCAACTGTTCACCTG 543
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuCysGluIleLysSer 140
DB 544 CTTGCGAGTGGGTTTCTAGTCTCTCGTGTTCCTCGGCTCTGTGAAGAAAATCC 603
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaIleuIleGlyProThr 160
DB 604 ACTGTGTCCTCACTGACATTTCTCGAGCCCTGCTTACCTGTTCCAAACATTTGGCCAAACC 663
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleuGlyLeu 180
DB 664 CGAATCTTCCCAATCTTATCTTGCTGCGACGAGATGCTCTCAACAGAGCTGATG 723
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProIleProAspPhe 200
DB 724 CAGCAGAAATGGGATTGGTTATGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 783
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
DB 784 ATCCCGAGCTCATTTCTCGCTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG 843
QY 221 ProTrpLeuAspLysSerValAspPheIleGluIleValAlaIleSerAsnGlyCysVal 240
```

Db 844 CCGTGGTGCACAAATCAGTATTCATTGAGAAAGCAAAAGCCCTCCCAATGATGTGT 903
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db 904 CTAGTGCACTGTTTACCTGGATCTCCCGCTCCGACCATGCTATGCGCTACATCATG 963
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLysArgProThr 280
Db 964 AAGAGCATGACATGCTTTAGATGAGCTTACAGATTGTGAAAAGAAAAGACCTTACT 1023
QY 281 ILeSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
Db 1024 ATATCTCCAAACTCAATTTCTGGGCCCAACTCTGACTATGAGAAAGATTAAAGAC 1083
QY 301 GlnThr 302
Db 1084 CAGACT 1089
RESULT 12
ABV20833
ID ABV20833 standard; cDNA; 5145 BP.
XX
AC ABV20833;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 20824.
XX
KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
PI
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3419; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 3,44e-175
Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Query Match: 99.23%
DB: 5
Gaps: 0
US-10-029-345a-109_copy_1_302 (1-302) x ABV20833 (1-5145)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
Db 589 ATGGCCCATAGATGATTTGGAACCTCAATTTGTTACTGAGAGGTTGGCTCTGCTGAA 648
QY 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db 649 AGTGAACCGAAGAAAGTGTGCTTAATTGATAGCCGCGCAATTGTGGAATACATCATCC 708
QY 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerLeuMetLysArgLeuGlnGln 60
Db 709 CACATTTTGAAGCCCATTAATATCACTGCTCCAACTTATGAAAGGATTGCAACAG 768
QY 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisLysValAspIleAsp 80
Db 769 GACAAAGTGTATTTACAGAGCTCATCCAGCAATTCCAGCAAAACATAGGTTGACATTGAT 828
QY 81 CysSerGlnLysValValValIleAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 829 TGCAGTCCAGAGGTTGTATGTTTACGATCAAACTCCCAAGATGTTCCTCTCTCTTCA 888
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu 120
Db 889 GACTGTTTCTCACTGACTCTCGGTAACCTGGAAGAGCTTCAACTCTTCACTG 948
QY 121 LeuAlaGlyIlePheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlyLysSer 140
Db 949 CTTCGAGGTGGGTTCGATGTTCTGCTGTTTCCCTGCGCTCTGTCGAAAGAAATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 1009 ACTCTAGTCCCTACCTGCACTTCTCAGCCTTCTTACCTGTTGCCAACAATGGGCCAAC 1068
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGlnLeuIle 180
Db 1069 CGAATTTCTCCCAATCTTTATCTTGCGTCCAGCGAAGATGCTCTCAAGAGGCTGATG 1128
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
Db 1129 CAGCAGAAATGGGATGTTATGTGTTAATGTCAGCAATVACTGTCCAAAGCTGACTTT 1188
QY 201 ILeProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
Db 1189 ATCCCGGAGTCTCATTTCTCGTGATGCTTGAAATGACAGCTTTTGGAGAAAATTTTG 1248
QY 221 ProTyrPheAspLysSerValAspPheIleGlnLysAlaLysAsnGlyCysVal 240
Db 1249 CCGTGGTTGGAACAATTCAGTATTTCTTGAAGAAAGAAAAGCTCCAAATGATGTGTT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db 1309 CTAGTGCACTGTTTACCTGGATCTCCCGCTCCGACCATGCTATGCTCATCATCATG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLysArgProThr 280
Db 1369 AAGAGCATGACATGCTTTAGATGAGCTTACAGATTGTGAAAAGAAAAGACCTACT 1428
QY 281 ILeSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
Db 1429 ATATCTCCAAACTCAATTTCTGGGCCCAACTCTGACTATGAGAAAGATTAAAGAC 1488
QY 301 GlnThr 302
|||||

Db 1489 CAGACT 1494

RESULT 13

ID ABV21080

ABV21080 standard; cDNA; 5145 BP.

XX

XX

AC ABV21080;

XX

DT 13-SEP-2002 (first entry)

XX

DE Human prostate expression marker cDNA 21071.

XX

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KM pharmacogenomic marker; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200160860-A2.

XX

PD 23-AUG-2001.

XX

PF 20-FEB-2001; 2001WO-US005171.

XX

PR 17-FEB-2000; 2000US-0183319P.

XX

PR 16-MAR-2000; 2000US-0189662P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

XX

XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX

PI Schlegel R, Endege WO, Monahan JE;

XX

DR WPI; 2001-662795/76.

XX

PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful

PT for detecting presence of prostate cancer, stage of prostate cancer.

XX

XX

PS Claim 1; Page 3481; 11750bp; English.

XX

XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the

CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX

SO Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 3,44e-175 Length: 5145

Score: 1552.00 Matches: 300

Percent Similarity: 99.67% Conservative: 1

Best Local Similarity: 99.34% Mismatches: 1

Query Match: 99.23% Indels: 0

DB: 5 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x ABV21080 (1-5145)

QY 1 MetAlhigIuWetIlegIYThrGlnIleValThrGlnArgIleValAlaIleuGln 20

DB 589 ATGGCCATGAGATGATGGAACCTCAATGTTACTGAGGCTTGCGCTGCGGAA 648

QY 21 SerGIYThrGIuIuValIleuIleuIleuSerArgProPheValGIuTYrAnThrSer 40

Db 649 AGTGAACGAAAAAGTGCTGCTAATTGATGACGGCCATTGTGGATATCAATACATCC 708

QY 41 HieIleuGlnUAlalleenIleancYserIleuMelIyArgrIleuGln 60

Db 709 CACATTTTGAAGCCATTATATCACTGCTCCMACCTTATGAGGAAGGTTGCAACG 768

QY 61 AapIyValIleuIleThrGIuIleuIlegInIIsSerAlaIyehIshIyValAapIleap 80

Db 769 GACAAAGTTTATTTATTCAGAGCTCATCCAGCTTCCAGCAACATTAAGTTGACATTGAT 828

QY 81 CySerGIuIyValIyValIyArpGlnSerSerGIuAapValAIsereIuSerSer 100

Db 829 TCGAGTCAGAAAGTTGTAGTTATACATCAAGCTCCCAAGATGTCCTCTCTCTCA 888

QY 101 AapCyPheIleuThrValIleuGlyIyIleuGlnIySerPheIenSerValIhIleu 120

Db 889 GACTGTTTCTCACTGACTTCTGGGTAACTGGAAGAGCTTCAACTGTTCACCTG 948

QY 121 LeuAlaGIyGIyPheAlaIupPheSerArgCyPheProGIyIleuCyGIuGIyIySer 140

Db 949 CTTCAGAGTGGGTTTGTCTAGATTCCTGCTGTTTCCCTGGCTCTGTGAAGAAATCC 1008

QY 141 ThrIleuValProThrCySIsereGlnProCyIleuProValAlaAnIleGIyProthr 160

Db 1009 ACTCTAGTCCCTACCTGATTTCTAGCTTCTGCTTCTTACCTTGCCAACTTGGGCCAAC 1068

QY 161 ArgIleIleuProAnIleuIyIleuGlyCyGIuIyArpValIleuAnIyIleuIle 180

Db 1069 CGAATTCCTCCCAATCTTATCTTGCTCCAGCGAGATGCTCTCAACAGAGCTGATG 1128

QY 181 GlnGIuAnGIyIleGIyTYrValIleuAnAlaSerTYrThCyPProIyPProAapPhe 200

Db 1129 CAGCAGAAATGGAGTTGTATGTGTTAAATCCAGCAAAATCTGTCCAAAGCTGACTTT 1188

QY 201 IleProIuSerIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 220

Db 1189 ATCCCGAGATCTCATTTCTGCGGTGCTGGAATGACAGCTTTGTGAGAAATTTTG 1248

QY 221 ProThrIleuAapIySerValIshPheIleGIuIyAlaIyAlaSerAnGIyCyVal 240

Db 1249 CCGTGTGTGACAAATCACTAGATTTTCATTTGAGAAAGCAAAAGCTTCAATGATGTGTT 1308

QY 241 LeuValhIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 260

Db 1309 CTAGTCACTGTTTACTGGAATCTCCCTCCGACACATGCTATCGCTACATCATG 1368

QY 261 LyArgMetAapMetSerIleuAapGIuAlaTYrArgPheValIyGIuIyArgProthr 280

Db 1369 AAGAGATGACATGCTTTAGTGAAGCTTACAGTTTGTGAAGAAAAAGACCTACT 1428

QY 281 IIsSerProAnPheAnPheIleuGIuIleuIleuAapTYrGIuIyIyIyIyIyIyIy 300

Db 1429 ATATCTCCCAACTTCAATTTTCTGGGCAACTCTGACTATGAGAAAGATTAAGAAC 1488

QY 301 GlnThr 302

Db 1489 CAGACT 1494

RESULT 14

ID ABV26680

ABV26680 standard; cDNA; 5145 BP.

XX

XX

AC ABV26680;

XX

DT 16-SEP-2002 (first entry)

XX

DE Human prostate expression marker cDNA 26671.

XX

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KM pharmacogenomic marker; gene; ss.

XX

OS Homo sapiens.

PN MO200160860-A2.
XX 23-AUG-2001.
PD
PF 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 5388-5389; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy
XX of progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX
XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 3,44e-175 Length: 5145
XX Score: 1552.00 Matches: 300
XX Percent Similarity: 99.67% Conservative: 1
XX Best Local Similarity: 99.34% Mismatches: 0
XX Query Match: 99.23% Indels: 0
XX DB: Gaps: 0
XX
XX US-10-029-345A-109_COPY_1_302 (1-302) X ABV26680 (1-5145)
XX
QY 1 MetAlahisgluWetllleglyThgnlleValThrgluArgleuValAlaleuLeuGlu 20
DB 589 ATGGCCCATGAGATGATGAACTCAATTGTGCTGAGAGGTTGGCTCGCTGGAA 648
QY 21 SerGlyThrgluValleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 649 AGTGAACGAGAAAGAGCTGCTGTAATGATGATGCGGCAATTTGGGAATCAATATAC 708
QY 41 HislleleuGluAlaleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 709 CACATTTGGAGACCATTAATATCAACGCTCCCAAGCTTAAGAGGAGGTTGCAACG 768
QY 61 AsplyValleuLeuThgGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 769 GACAAAGCTTAATACAGAGCTCATCCAGCATTCAGGAAACATTAAGTTGACATTGAT 828
QY 81 CysSerGlnLysValValValValValValValValValValValValValVal 100
DB 829 TCGAGTCAGAGGTTTACATTACATCAAGCTCCCAAGAGTGGCTCTCTCTCTTCA 888
QY 101 AspCysPheLeuThrValleuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu 120
DB 889 GACTGTTTCTCACTGACTTCTGGGTAACTGGAGAGAGCTTCACTCTTCACTG 948

QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
DB 949 CTTCAGAGTGGGTTTCTGAGTTCTCTGTTTCTCTGCTGCTGCTGAGAGAAATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnLleGlyProThr 160
DB 1009 ACTTATGTCCTTACCTGATTTCTGAGCTTGGCTTACCTGTTGCCAACTTGAGCCAAAC 1068
QY 161 ArgIleLeuProAsnLeuLysLysGlyCysGlnArgAspValLeuAsnLysGluLeu 180
DB 1069 CGAATTCCTCCCAATCTTATCTTGGCTGACAGAGATGCTCTCAACAAAGAGCTGATG 1128
QY 181 GlnGlnAsnGlyLleGlyValValleuAsnLleSerValThrCysProLysProAspPhe 200
DB 1129 CAGCAGAAATGGATTTGTTTATGTTTAAATGACAGCAATACCTGTCCAAGCCTGCTT 1188
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
DB 1189 ATCCCGAGTCTCATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
QY 221 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
DB 1249 CCGTGTTGACAAATTCAGTATTTTCAATTCAGAAACAAAGCTTCATGATGATGTTG 1308
QY 241 LeuValHisCysLeuAlaGlyLleSerArgSerAlaThrIleAlaIleAlaValIleMet 260
DB 1309 CTAGTGCACTGTTTGTGCTGGATCTCCGCTCCGCAACATTCCTATCGCTCATCATG 1368
QY 261 LysAlaMetAspMetSerLeuAspGluAlaValArgPheValLysGluValArgProThr 280
DB 1369 AAGAGGATGACATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1428
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspValArgLysLysIleLysAsn 300
DB 1429 ATATCTCCAACTTCAATTTCTGGGCCAATCTGAGCTAGTAAAGAAAGATTTAAGAC 1488
QY 301 GlnThr 302
DB 1489 CAGACT 1494
XX
XX
XX RESULT 15
XX ABV20978
XX ID ABV20978 standard; cDNA; 5145 BP.
XX AC ABV20978;
XX
XX 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 20969.
XX
XX DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX PN WO200160860-A2.
XX
XX 23-AUG-2001.
XX PD
XX PF 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX

DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 3451; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	3,44e-175	Length:	5145
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-109_COPY_1_302 (1-302) x ABV20978 (1-5145)

QY 1 MetAlhiSGluMetileGlyThrGlnileValThrGluArgLeuValAlaLeuLeuGlu 20
DB 589 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 648
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 649 AGTGAACGGAAAGAGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 708
QY 41 HisileLeuGluValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 709 CACATTTGGAAACCATTAATATCACTGCTCCAGCTTATGAAGCAAGTTGCAACAG 768
QY 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 769 GACAAAGTTTAAATACAGAGCTATCCAGCATTCAGCAACATTAAGTTGACATTGAT 828
QY 81 CysSerGlnIleValValValValValValValValValValValValValValVal 100
DB 829 TGCAGTCAGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
DB 889 GACTGTTTCTCACTGACTTCTGGGTAACCTGGAAGAGCTTCAACTGTTCACTG 948
QY 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuGlyLysSer 140
DB 949 CTGGCAGGTGGTTGGTGGATTCCTCGTTGTTCCCTGGCTCTGTGAAGAAATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnileGlyProThr 160
DB 1009 ACTCTAGTCCCTACCTGCACTTCTCAGCTTCTCAGCTTCTCAGCTTCTCAGCTTCT 1068
QY 161 ArgIleLeuProAsnLeuThrLysLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
DB 1069 CGAATCTTCCCAATCTTATCTTGGCTGCCAGCGAGATCTCTCAACAAAGAGCTGATG 1128
QY 181 GlnGlnAsnGlyIleGlyThrValLeuAsnAlaSerThrCysProLysProAspPhe 200
DB 1129 CACGAGATGGATGGTGTATGTTAAATGACAGCAATACCTGTCCAAAGCTGACTTT 1188

QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
DB 1189 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
QY 221 ProThrLeuAspLysSerValAspPheileGlyLysAlaLysAlaSerAsnGlyCysVal 240
DB 1249 CCGTGGTTGACAAATACAGTATTCATTGAGAAAGCAAAAGCTTCAATGATGATGTT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleValIleMet 260
DB 1309 CTAGTCCACTGTTTACCTGGAGATCTCCGCTCCGCAACATCCCTATCGCTTACATCATG 1368
QY 261 LysArgMetAspSerLeuAspGluAlaThrArgPheValLysGluLysArgProThr 280
DB 1369 AAGAGATGACATGCTTTTAAATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1428
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspThrGlyLysIleLysAsn 300
DB 1429 ATATCTCCAAACTTCAATTTTCTGGGCCCAACTCTGAGCTATGAGAAAGATTAAAGAC 1488
QY 301 GlnThr 302
DB 1489 CAGACT 1494

Search completed: June 21, 2004, 21:33:10
Job time : 347.168 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 21:18:06 ; Search time 372.269 Seconds

(without alignments)
3716.370 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302

Perfect score: 1564

Sequence: 1 MAHEMGTQVTRERLVALE.....PNEFGLQDLRYEKIKINQI 302

Scoring table:
BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xjh
-Q=/cgn2_1/USPTO.epool/US10029345/rnat 21062004 164146 29071/app.query.fasta_1.1294
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rmpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human4.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10029345 @CGN_1_1_33_@rnat 21062004 164146 29071
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications_NA:*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	1552	99.2	1998	9	US-09-816-494-3	Sequence 3, Appl1
2	1552	99.2	1998	13	US-10-377-072-27	Sequence 27, Appl1
3	1552	99.2	2071	16	US-10-072-012-257	Sequence 257, App
4	1552	99.2	2732	13	US-10-168-506-2	Sequence 2, Appl1
5	1552	99.2	3059	17	US-10-257-026-1	Sequence 1, Appl1
6	1552	99.2	3496	9	US-09-964-277-1	Sequence 1, Appl1
7	1552	99.2	3544	9	US-09-816-494-1	Sequence 1, Appl1
8	1552	99.2	3544	16	US-10-377-072-25	Sequence 25, Appl1
9	1552	99.2	3625	13	US-10-425-114-26234	Sequence 26234, A
10	1552	99.2	3766	13	US-10-343-357-17	Sequence 17, Appl1
11	1552	99.2	4790	16	US-10-648-593-115	Sequence 115, App
12	1545	98.8	2102	17	US-10-094-459-673	Sequence 673, App
13	1544	98.7	2966	13	US-10-296-115-520	Sequence 520, App
14	1504.5	96.2	2200	13	US-10-072-012-255	Sequence 255, App
15	1194.5	76.4	3332	9	US-09-964-277-20	Sequence 20, Appl1
16	985	63.0	2453	14	US-10-005-858-1	Sequence 1, Appl1
17	943	60.3	2476	13	US-10-220-120-28	Sequence 28, Appl1
18	837	53.5	2039	13	US-10-072-012-265	Sequence 265, App
19	469	30.0	1337	9	US-09-964-899-46	Sequence 46, Appl1
20	469	30.0	1449	12	US-10-052-482-192	Sequence 192, App
21	469	30.0	1830	15	US-10-346-356-1	Sequence 1, Appl1
22	469	30.0	1909	12	US-10-052-482-191	Sequence 191, App
23	469	30.0	1909	13	US-10-058-270A-129	Sequence 129, App
24	468	29.9	1452	12	US-10-052-482-189	Sequence 189, App
25	468	29.9	1977	12	US-10-052-482-188	Sequence 188, App
26	467.5	29.9	2109	17	US-10-305-720-1135	Sequence 1135, Ap
27	467.5	29.9	2109	16	US-10-641-643-946	Sequence 946, App
28	467.5	29.9	2390	13	US-10-342-887-698	Sequence 698, App
29	467.5	29.9	2390	13	US-10-172-118-698	Sequence 698, App
30	467.5	29.9	2390	15	US-10-175-523-183	Sequence 183, App
31	467.5	29.9	2390	16	US-10-199-221-4	Sequence 4, Appl1
32	467.5	29.9	2390	17	US-10-117-597-4	Sequence 4, Appl1
33	467.5	29.9	2649	16	US-10-133-937-48	Sequence 48, Appl1
34	467.5	29.9	2649	16	US-10-159-563-48	Sequence 48, Appl1
35	467.5	29.9	2792	10	US-09-971-329-69	Sequence 69, Appl1
40	461.5	29.5	2104	16	US-10-388-934-654	Sequence 654, App
41	461.5	29.5	2104	16	US-10-191-803-270	Sequence 270, App
42	461	29.5	3286	10	US-09-971-392-9	Sequence 9, Appl1
43	426	27.2	877	9	US-09-764-853-158	Sequence 158, App
44	426	27.2	2240	16	US-10-305-720-1100	Sequence 1100, Ap
45	426	27.2	2241	16	US-10-174-460-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-816-494-3

Alignment Scores: 1.89e-201 Length: 1998
Pred. No.: 1.89e-201

Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Query Match: 99.23%
DB: 9

Matches: 300
Conservative: 1
Mismatch: 1
Indels: 0
Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-816-494-3 (1-1998)

```
QY 1 MetAlaHISgluMetilegIYThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 1 ATGGGCCATGAGATGATGAGAACTCAAAATGTTACTGAGAGGTGGCTCTCTCGAA 60
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluIYrAsnThrSer 40
DB 61 AGTGAACGGAAAAAGTGTGCTGATGATAGCCGGCATTTGGGAATCAATCAATCC 120
QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleuMetIYsArgLeuGln 60
DB 121 CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCACAG 180
QY 61 AspIYsValIleuIleThrGluLeuIleGlnHisSerAlaYsHisIYsValAsp 80
DB 181 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCACGAAACATAGGTTGACATGAT 240
QY 81 CysSerGlnYsValIValIYrAspGlnSerSerGlnAspValAlaSerLeuSer 100
DB 241 TGCAGTCAAGAGGTGATGTTAGATCAAGTCCCAAGCTTATGAAAGTGGTCAACAG 300
QY 101 AspCysPheLeuThrValLeuLeuGluYsIleGlnIYsSerPheAsnSerValHisIleu 120
DB 301 GACGTTTTCATCTGATCTTGGGTAACGTGAGAGAGCTTCAACTCTGTTCACTG 360
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyIYsSer 140
DB 361 CTTCAGAGTGGTGTGCTGAGTTCCTCTGTTGTTCCCTGCGCTCTGTAAGAGAAATTC 420
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 421 ACTCTAGTCCCTACCTGCACTTCTGACCTGCTTCACTGCTGCAACATTTGGCAACC 480
QY 161 ArgIleLeuProAsnLeuIYrLeuGlyCysGlnArgAspValLeuAsnIYsGluLeuIle 180
DB 481 CGAATTCCTCCCAATCTTATCTTGCTCCAGCGAGATGCTCTCAACAAGAGCTGATG 540
QY 181 GlnGlnAsnGlyIleGlyIYrValLeuAsnAlaSerIYrThrCysProIYsProAspPhe 200
DB 541 CAGCAGAAATGGAGATTTGTTATGTTAAATGCCAGCAATACCTGTCCAAAAGCTGACCTT 600
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlyIYsIleLeu 220
DB 601 ATCCCGAGCTCATTTCTCGCTGCTGCTGCAAGTCTCTCAACAAGAGCTGATG 660
QY 221 ProTrpLeuAspIYsSerValAspPheIleGlyIYsAlaYsAlaSerAsnGlyCysVal 240
DB 661 CCGTGTTCGACAAATCCAGTATGATTCATGGAAGCAAAAGCTCCAAATGATGTT 720
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIYrIleMet 260
DB 721 CTAGTGCACTGTTAGTGGGATCTCCGCTCCGCAACATCGCTATACGCTATCATG 780
QY 261 LysArgMetAspMetSerLeuAspGluAlaIYrArgPheValIYsGlyIYsArgProThr 280
DB 781 AAGAGAGTGAACATGCTTTAGATGAACTTACAGATTTGGAAGAAAAAGACTACT 840
QY 281 IleSerProAsnPheAsnPheLeuGlyIYsIleuAsnAspIYrGlyIYsIYsIleYsAsn 300
DB 841 ATATCTCAAACTCAATTTCTGGGCCAATCTCTGACATATGAGAAAGATTAAAGAAC 900
QY 301 GlnThr 302
DB 901 CAGACT 906
```

RESULT 2

```
US-10-377-072-27
; Sequence 27, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Myoung
; APPLICANT: Tsai, Feng-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: MP103-0180NM1M
; CURRENT APPLICATION NUMBER: US/10/377,072
; CURRENT FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; PRIOR FILING DATE: 2001-03-22
; Remaining prior application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1998)
; US-10-377-072-27

Alignment Scores:
Pred. No.: 1,896-201
Score: 1552.00 Length: 1998
Percent Similarity: 99.67% Matches: 300
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 16 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x US-10-377-072-27 (1-1998)
QY 1 MetAlaHISgluMetilegIYThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 1 ATGGGCCATGAGATGATGAGAACTCAAAATGTTACTGAGAGGTGGCTCTCTCGAA 60
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluIYrAsnThrSer 40
DB 61 AGTGAACGGAAAAAGTGTGCTGATGATAGCCGGCATTTGGGAATCAATCAATCC 120
QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleuMetIYsArgLeuGln 60
DB 121 CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCACAG 180
```



```
Qy 61 AsplvleValleuileThrgluLeuileglnHiseSerAlaIysHsiIyValIAspIleAsp 80
Db 181 GACAAAGTGTAAATTACAGAGCTCATTCACGAAACCTAAGTTGACATTGAT 240
Qy 81 CysSerGlnIyValValIYrAspGlnSerSerGlnAspValAlaSerLeuSer 100
Db 241 TGCAGTCAGAAAGTTGATTAGATCAAGCAAGCTCCCAAGATGTGCTCTCTTCA 300
Qy 101 AspCyPheLeuThrValLeuLeuGlyLyLeuGlnIySerPheAnsSerValHleu 120
Db 301 GACTGTTTCTCACTGACTTCTGGGTAAACTGGAGAAGCTTCACTGTTCACCTG 360
Qy 121 LeuAlGlyLyPheAlaGluPheSerArgCyPheProGlyLeuCyGlnGlyLySer 140
Db 361 CTTCAGAGTGGTTTCTGAGTTCTCTGTTGTTTCCCTGCTCTTGAAGAAATCC 420
Qy 141 ThrLeuValProThrCySileSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
Db 421 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGCTTACCTGTTCACATTTGGGCCAAC 480
Qy 161 ArgIleLeuProAnsLeuIYrLeuGlyCyGlnArgAspValLeuAnsIyGlnLeuIle 180
Db 481 CGAATCTCTCCCAATCTTATCTTGGCTGCAGGAGATGTCTCAACAAGAGCTGATG 540
Qy 181 GlnGlnAsnGlyIleGlyYrValLeuAsnAlaSerYrThrCySProLyProAspPhe 200
Db 541 CAGAGAGATGGATGGTATGTTAAATGACAGCAATCTGTCCAAAGCTGACTTT 600
Qy 201 IleProGlnSerHsiPheLeuArgValProValAsnAspSerPheCyGlnIyLeu 220
Db 601 ATCCCGAGCTCATTTCCGCGGTGCTCGTGAATACAGCTTTTGAGAAATTTTG 660
Qy 221 ProTrpLeuAspLySerValAspPheIleGlnIyAlaIyAlaSerAnsIyCyVal 240
Db 661 CCGGTGTGAGCAAAATCATGATTTCTATTGAGAAACCAAAAGCTCCAAATGATGTT 720
Qy 241 LeuValHsiCySerLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaYrIleMet 260
Db 721 CTAGTGACATGTTAGTGGGATTCCTCCGCTCCGACCAATCGCTATCGCTTACATATG 780
Qy 261 LysArgMetAspMetSerLeuAspGlnAlaIYrArgPheValIyGlnIySArgProThr 280
Db 781 AAGAGATGAGATGCTCTTTAGATGAGAGCTTACAGATTTTGAAAGAAAAAGCTTACT 840
Qy 281 IleSerProAnsPheAnsPheLeuGlyGlnLeuLeuAspYrGlnIySlyIleIyAsn 300
Db 841 ATATCTCCAACCTTCAATTTTCTGGCCCACTCCTGAGATATGAGAAAGATTAAAGAC 900
Qy 301 GlnThr 302
Db 901 CAGACT 906
```

RESULT 3
US-10-072-012-257

```
; Sequence 257, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernov, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zetzsche, Bryan
; APPLICANT: Patursajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Bsha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taudier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
```

```
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072, 012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265, 102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265, 514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265, 517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265, 412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265, 395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266, 406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266, 767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267, 057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266, 975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267, 459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 257
; LENGTH: 2071
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-072-012-257

Alignment Scores:
Pred. No.: 2, 01e-201 Length: 2071
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
Gaps: 0
DB: 13

US-10-029-345A-109_COPY_1_302 (1-302) x US-10-072-012-257 (1-2071)

Qy 1 MetAlaHsiGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
Db 61 ATGGCCCATGAGATGATGGAGCTCAATGTACTGAGAGGTTGGGCTGCTGGAA 120
Qy 21 SerGlyThrGlnIyValLeuLeuIleAspSerArgProPheValGluYrAsnThrSer 40
Db 121 AGTGAACGAAAGAGTGTCTGCTAATTGATGACCGGCCCATTTGTGGAATCATACATCC 180
Qy 41 HsiIleGlnIyAlaIleAsnIleAnsCySerIyLeuMetIySArgLeuGln 60
Db 181 CACATTTTGAACCATATATATCACTGCTCCAAAGCTTATGAAAGCAAGTTGCAACAG 240
Qy 61 AsplvleValleuileThrgluLeuileglnHiseSerAlaIysHsiIyValIAspIleAsp 80
Db 241 GACAAAGTGTAAATTACAGAGCTCATTCACGAAACCTAAGTTGACATTGAT 300
Qy 81 CysSerGlnIyValValIYrAspGlnSerSerGlnAspValAlaSerLeuSer 100
Db 301 TGCAGTCAGAAAGTTGATTAGATCAAGCAAGCTCCCAAGATGTGCTCTCTTCA 360
Qy 101 AspCyPheLeuThrValLeuLeuGlyLyLeuGlnIySerPheAnsSerValHleu 120
Db 361 GACTGTTTCTCACTGACTTCTGGGTAAACTGGAGAAGCTTCACTGTTCACCTG 420
```

```

QY      121  LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
Db      421  CTGGACAGTGGGTTGGTGTGAGTTCCTCGTTGTTCCCTGGCCTCTGTGAAAGAAATCC 480
QY      141  ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db      481  ACTCTAGTCCCTACCTGCAGATTCTCGACCTTGGCTTAACCTGTGCCAACAATGGGCGCAAC 540
QY      161  ArgIleLeuProAsnLeuTyrIleuGlyCysGlnArgPheValLeuAsnIleGlyLeu 180
Db      541  CGAATCTTCCCAATCTTATCTGTGCTGCCAGCAGATGTCCTCAACAAGAGCTGATG 600
QY      181  GlnGlnAsnGlyIleGlyTyrValLeuAsnIleSerTyrThrCysProLysProAspPhe 200
Db      601  CAGCAGAAATGGGTTGGTGTATGTGTAATGCCAGCAATACCTGTCCAAAGCCGACCTTT 660
QY      201  IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
Db      661  ATCCCGAGTCTCATTTTCTGCTGCTGCTGGAATGACAGCTTTTGTGAAATTTTG 720
QY      221  ProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
Db      721  CCGTGGTTGGACAATCAGTACGATTTCTTATGAGAAAGCAAAAGCCTCCAAATGATGTGT 780
QY      241  LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db      781  CTAGTGCACTGTTTAGCTGGGATCTCCGCTCCGCCACATCGCTATCGCTACATCATG 840
QY      261  LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGlyLysArgProThr 280
Db      841  AAGAGGATGGACATCTCTTTAGATGAGCTTACGATTTGTGAAAGAAAGAAAGAAAGCTTACT 900
QY      281  IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
Db      901  ATATCTCCAAACTTCATTTTCTGGGCCAATCTCTGGAATATGAGAAAGATTAAAGAAC 960
QY      301  GlnThr 302
Db      961  CAGACT 966

```

RESULT 4

```

US-10-168-506-2
; Sequence 2, Application US/10168506
; Publication No. US20040053229A1
; GENERAL INFORMATION:
; APPLICANT: PLOMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: HILL, RON
; APPLICANT: PLANAGAN, PETER
; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
; FILE REFERENCE: 038602/1351
; CURRENT APPLICATION NUMBER: US/10/168,506
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: PCT/US00/34736
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-168-506-2

```

Alignment Scores:

```

Pred. No.: 3,24e-201
Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Query Match: 99.23%

```

```

Length: 2732
Matches: 300
Conservative: 1
Mismatches: 1
Indels: 0

```

```

DB: 13 Gaps: 0
US-10-029-345a-109_copy_1_302 (1-302) x US-10-168-506-2 (1-2732)
QY      1  MetAlaHisGlyMetIleGlyThrGlnIleValIleThrGluArgLeuValAlaLeuLeuGlu 20
Db      538  ATGGCCCATGAGATGATGATGGAACTCAAAATGTACTAGAGAGGTTGGCTCTCTCGAA 597
QY      21  SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db      598  AATGGAAACGCAAAAAGTCTGCTAATGATGACCGGCACTTTGGAAATCAATACATCC 657
QY      41  HisIleLeuGluAlaIleAsnIleAsnCysSerTyrSleuMetLysArgLeuGln 60
Db      658  CACATTTTGGAAAGCATTAAATATCACTGCTCCAAAGCTTATGAAAGCGAAGTTGCAACG 717
QY      61  AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db      718  GACAAAGTTAAATTACAGAGCTCATCCAGCATTCAGCAAACTAAAGTTGACATTGAT 777
QY      81  CysSerGlnLysValValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db      778  TGCACTGAGAAAGTTGATGATTAACGATCAAGCTCCCAAGATGTTCCTCTCTCTCA 837
QY      101  AspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu 120
Db      838  GACTGTTTCTCACTGATCTTCTGGGTAACCTGGAAGAGCTTCAACTCTGTCACTG 897
QY      121  LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
Db      898  CTTCAGAGTGGGTTGGTGTGAGTTCCTCGTTGTTCCCTGGCCTCTGTGAAAGAAATCC 957
QY      141  ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db      958  ACTCTAGTCCCTACCTGCAGATTCTCAGCCTTCTTAACCTGTTGCCAAATGGGCGCAACC 1017
QY      161  ArgIleLeuProAsnLeuTyrIleuGlyCysGlnArgPheValLeuAsnIleGlyLeu 180
Db      1018  CGAATCTTCCCAATCTTATCTGTGCTGCCAGCAGATGTCCTCAACAAGAGCTGATG 1077
QY      181  GlnGlnAsnGlyIleGlyTyrValLeuAsnIleSerTyrThrCysProLysProAspPhe 200
Db      1078  CAGCAGAAATGGGTTGGTGTATGTGTAATGCCAGCAATACCTGTCCAAAGCCGACCTT 1137
QY      201  IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
Db      1138  ATCCCGAGTCTCATTTTCTGCTGCTGCTGGAATGACAGCTTTTGTGAAATTTTG 1197
QY      221  ProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
Db      1198  CCGTGGTTGGACAATCAGTACGATTTCTTATGAGAAAGCAAAAGCCTCAATGATGTGT 1257
QY      241  LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db      1258  CTAGTGCACTGTTTAGCTGGGATCTCCGCTCCGCCACATCGCTATCGCTACATCATG 1317
QY      261  LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGlyLysArgProThr 280
Db      1318  AAGAGGATGGACATCTCTTTAGATGAGCTTACGATTTGTGAAAGAAAGAAAGAAAGCTTACT 1377
QY      281  IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
Db      1378  ATATCTCCAAACTTCATTTTCTGGGCCAATCTCTGGAATATGAGAAAGATTAAAGAAC 1437
QY      301  GlnThr 302
Db      1438  CAGACT 1443

```

RESULT 5

```

US-10-257-026-1
; Sequence 1, Application US/10257026
; Publication No. US20040086859A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10X.DMS
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3059
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)..(2121)
US-10-257-026-1

Alignment Scores:
Pred. No.: 3,93e-201 Length: 3059
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
Gaps: 0
17

US-10-029-345A-109_COPY_1_302 (1-302) x US-10-257-026-1 (1-3059)

QY 1 MetAlAHISGLWETILEGIYThRGInILEValThRGUARGLeuValAlaleuLeuGln 20
Db 127 ATGGCCCATAGATGATGTGAACTCAATGTTACTGAGAGGTGGGCTCGCTGGAA 186
QY 21 SerGIYThRGUlyValleuLeuILEAspSerARpPheValGluTYrAenThSer 40
Db 187 AGTGGAAAGGAAAGAGTGGCTTATGATAGCCGCCATTGTGGANTCAATACATCC 246
QY 41 HisILEuGlnAlaleuLeuILEAsnCySerIlyleuWetIySARgLeuGln 60
Db 247 CACATTTGGAGCCATTATATCACTGCTCCAAAGTTTGAAGCAAGTTGCAACAG 306
QY 61 AspIlyValleuLeuThRGUleuILEGlnHISerAlalySHIlySValAspIleAsp 80
Db 307 GACAAAGTGTATATTCAGAGCTCATCCAGCATTCAGCGAAACATAGAGTTGACATTGAT 366
QY 81 CySerGlnIlyValIValIYrAspGlnSerSerGlnAspValAlaserLeuSerSer 100
Db 367 TGCAGTCAAGAGTGTGATTTACGATCAAGCTCCCAAGATGTGGCTCTCTCTTCA 426
QY 101 AspCyPheLeuThrValIleuLeuGlyLySLeuGlyIlySerPheAsnSerValIHISleu 120
Db 427 GACTGTTTCTCACTGTACTTCTGGGTAACCTGGAGAAAGCTTCAACTGTTCACTG 486
QY 121 LeuAlAGlyGlyPheAlAGlyPheSerArgCyPheProGlyLeuCySGluGlyLySer 140
Db 487 CTTCGAGGTGGGTTGCTGAGTCTCTCGTTGTTCCCTGCTGTGAGGAAATCC 546
QY 141 ThrLeuValProThrCySIIeserGlnProCyLeuProValAlaleuILEGlyProThr 160
Db 547 ACTCTGACTCTTACCTGCACTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCCAAAC 606
QY 161 ArgIleuProAsnLeuTYrLeuGlyCySGlnARgAspValIleuAsnIySGluLeu 180
Db 607 CGAATCTTCCCAATCTTTATCTTGGTGCAGAGAGATGCTCTCAACAGAGAGCTGATG 666
QY 181 GlnGlnAsnGlyIleGIYTYrValleuAsnAlaserTYrThCySPolySProAspPhe 200
Db 667 CACACAGATGGAGTGTGTATGTGTTAAATGCAAGCAATACCTGTCCAAAGCTGTACTTT 726
QY 201 IISerProAsnPheAsnPheLeuARGValProValAsnAspSerPheCySGluIySleu 220
Db 727 ATCCCGCAGTCTATTTCTCGCTGCTGCTGTGATATACAGCTTTGTGGAATAATTTG 786
QY 221 ProTIPheAspIlySerValAspPheILEGlyIySValIySAlaserAnGlyCyVal 240
Db 787 CCGTGGTGGACAAATACATGATTTTCATTGAGAAAGCAAAAGCTCTCAATGATGTGT 846
```

```

QY 241 LeuValHISCyLeuAlAGlyIISerARgSerAlaThrIleAlaleuIleMet 260
Db 847 CTAGTGACTGTGTAGCTGGAGATCTCCGCTCCGCAACATCGCTTACCTACATAG 906
QY 261 IySArgMetAspMetSerLeuAspGluAlalyARgPheValIySGlyIySARgProThr 280
Db 907 AAGAGATGGACATGCTTATAGATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 966
QY 281 IISerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTYrGluIySValIySAsn 300
Db 967 ATATCTCCAAACTTCAATTTCTGGGCAACTCTGACTATGAGAAAGATTAAGAAC 1026
QY 301 GlnThr 302
Db 1027 CAGACT 1032

RESULT 6
US-09-964-277-1
; Sequence 1, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-1

Alignment Scores:
Pred. No.: 4,94e-201 Length: 3496
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
Gaps: 0
9

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-964-277-1 (1-3496)

QY 1 MetAlAHISGLWETILEGIYThRGInILEValThRGUARGLeuValAlaleuLeuGln 20
Db 562 ATGGCCCATAGATGATGTGAACTCAATGTTACTGAGAGGTGGGCTCGCTGGAA 621
QY 21 SerGIYThRGUlyValleuLeuILEAspSerARpPheValGluTYrAenThSer 40
Db 622 AGTGGAAAGGAAAGAGTGGCTTATGATAGCCGCCATTGTGGANTCAATACATCC 681
QY 41 HisILEuGlnAlaleuLeuILEAsnCySerIlyleuWetIySARgLeuGln 60
Db 682 CACATTTGGAGCCATTATATCACTGCTCCAAAGTTTGAAGCAAGTTTCTCTCTTCA 741
QY 61 AspIlyValleuLeuThRGUleuILEGlnHISerAlalySHIlySValAspIleAsp 80
Db 742 GACAAAGTGTATATTCAGAGCTCATCCAGCATTCAGCGAAACATAGAGTTGACATTGAT 801
QY 81 CySerGlnIlyValIValIYrAspGlnSerSerGlnAspValAlaserLeuSerSer 100
Db 802 TGCAGTCAAGAGTGTGATTTACGATCAAGCTCCCAAGATGTTCCTCTCTCTTCA 861
QY 101 AspCyPheLeuThrValIleuLeuGlyLySLeuGlyIlySerPheAsnSerValIHISleu 120
Db 862 GACTGTTTCTCACTGTACTTCTGGGTAACCTGGAGAAAGCTTCAACTGTTCACTG 921
QY 121 LeuAlAGlyGlyPheAlAGlyPheSerArgCyPheProGlyLeuCySGluGlyLySer 140
Db 922 CTTCGAGGTGGGTTTCTGAGTCTCTCGTTGTTTCCCTGCTGTGAGGAAATCC 981
```

QY 141 ThrleuVal1ProthrCys11eserGlnProCysleuProVal1A1asn11eglyProthr 160
DB 982 ACTGAGTCCCTACTGATTTCTTACGCTTGCTTACTGCTTGTGCAACATTTGGGCAACC 1041
QY 161 Arg11leuProasnleuTyrlleuglyCysglnArgPheVal1leuasnlygluleu11e 180
DB 1042 CGAATTCCTCCCAATCTTATCTTGCTGCGCAGCAGATGTCTCAACAGAGCTGTATG 1101
QY 181 GlnGlnasnGly11eglyTyrlVal1leuasn11aserTyrlThrCysProlyProasphe 200
DB 1102 CAGCAGATGGAGTTAGTTAGTTAATGCAAGCAATACCTGTGCTCAACGCTGACTTT 1161
QY 201 IleProGlnSerHisPheleuArgVal1ProVal1asnSerPheCysglulys11leu 220
DB 1162 ATCCCGAGTCTCATTTCTGCGTGTGCTGATGATGACAGCTTTTGTGAGAAATTTTG 1221
QY 221 ProTrpleuAspLysSerVal1aspPhe11eglylysal1alyal1aserasnGlyCysVal 240
DB 1222 CCGGTGTGACAAATCAGTATGATTTCAATTGAGAAAGCAACCTCCAAATGATGTGT 1281
QY 241 LeuVal1HisCysleuVal1agly11eserArgSer11aThr11eal1a1aTyrl11eMet 260
DB 1282 CTAGTGACAGTGTAGTGTGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 1341
QY 261 LysArgMetAspMetSerleuAspGlu1a1aTyrlArgPheVal1lysglulysArgProthr 280
DB 1342 AAGAGGATGACATGTCTTTAGATGACATTACAGATTTGTGAAAGAAAGAAAGCTTACT 1401
QY 281 IleSerProasnPheasnPheleuglyGlnleuAspTyrlGlylylysal1elysasn 300
DB 1402 ATATCTCCAAATCTTATCTTGCGGCAACTCTGAGCTATGAGAAAGATTAAAGAAC 1461
QY 301 GlnThr 302
DB 1462 CAGACT 1467

RESULT 7

US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)...(2583)
US-09-816-494-1

Alignment Scores:

Pred. No.: 5,06e-201 Length: 3544
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-816-494-1 (1-3544)

QY 1 MetalaHisGluMet11Gly1ThrGln1leVal1ThrGluArgleuVal1A1leu1leuGlu 20
|||||

DB 589 ATGGCCCATGAGATGATTGGAACTCAAATTTGTTACTGAGAGCTTGCGTCTCTGGAA 648
QY 21 SerGlyThrGlu1lysal1leu1leu1easpSerArgProPheVal1GlyTyrlAsnThrSer 40
DB 649 AGTGAAAGGAAAAGAGCTGCTTATTTGATAGCGGCACTTTGTGGAATACATATCATCC 708
QY 41 His11leuGlu1a11easn11easnCysSerLysleuMetLysArgGleuGln 60
DB 709 CACATTTGGAAACCAATTAATATCACTGCTCAAGCTTATGAAAGCAAGGTTCACAG 768
QY 61 AspLysVal1leu1leu1ThrGlu1leu11eglnHisSer11a1yshi1ylysal1asp 80
DB 769 GACAAAGTTAATTAACAGAGCTATCCAGCATTCAGCAAAACATTAAGTTGACATTTGAT 828
QY 81 CysSerGln1ySal1Val1TyrlAspGlnSerSerGlnAspVal1aserleuSerSer 100
DB 829 TGCACTGAGAGGTGATGATTAAGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 888
QY 101 AspCysPheleuThrVal1leu1leuGlyLysleuGluLysSerPheasnSerVal1Hisleu 120
DB 889 GACTGTTTCTCACTGTAATCTTGCGTAACTGAGAAAGCTTCAACTGTTCACCTG 948
QY 121 LeuVal1agly1Phe11agluPheSerArgCysPheProGlyleuCysglulys11ySer 140
DB 949 CTGAGGTGGGTGTGCTGAGTCTCTCGTTGTTCCCTGGCTCTGTGAAAGAAATCC 1008
QY 141 ThrleuVal1ProthrCys11eserGlnProCysleuProVal1A1asn11eglyProthr 160
DB 1009 ACTGTAGTCCCTACCTGACATTTCTCAGCCCTGTGCTTACTGTTGCCAATTTGGGCCAAC 1068
QY 161 Arg11leuProasnleuTyrlleuglyCysglnArgPheVal1leuasnlygluleu11e 180
DB 1069 CGAATTCCTCCCAATCTTATCTTGCGTCCAGCAGATGTCTCCAAAGAGAGCTGATG 1128
QY 181 GlnGlnasnGly11eglyTyrlVal1leuasn11aserTyrlThrCysProlyProasphe 200
DB 1129 CAGCAGATGGAGTTAGTTAGTTAATGCAAGCAATACCTGTCCAAAGCTGACTTT 1188
QY 201 IleProGlnSerHisPheleuArgVal1ProVal1asnSerPheCysglulys11leu 220
DB 1189 ATCCCGAGTCTCATTTCTGCGTGTGCTGATGATGACAGCTTTTGTGAGAAATTTTG 1248
QY 221 ProTrpleuAspLysSerVal1aspPhe11eglylysal1alyal1aserasnGlyCysVal 240
DB 1249 CCGGTGTGACAAATCAGTATGATTTCAATTGAGAAAGCAACCTCCCAATGATGTGT 1308
QY 241 LeuVal1HisCysleu11agly11eserArgSer11aThr11eal1a1aTyrl11eMet 260
DB 1309 CTAGTGACAGTGTAGTGTGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 1368
QY 261 LysArgMetAspMetSerleuAspGlu1a1aTyrlArgPheVal1lysglulysArgProthr 280
DB 1369 AAGAGATGACATGTCTTTAGATGACATTACAGATTTGTGAAAGAAAGAAAGCTTACT 1428
QY 281 IleSerProasnPheasnPheleuglyGlnleuAspTyrlGlylylysal1elysasn 300
DB 1429 ATATCTCCAAATCTTATCTTGCGGCAACTCTGAGCTATGAGAAAGATTAAAGAAC 1488
QY 301 GlnThr 302
DB 1489 CAGACT 1494

RESULT 8

US-10-377-072-25
; Sequence 25, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Gluckmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.

APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Miyoung
TITLE OF INVENTION: Novel 25869, 25934, 26335, 50365, 21117,
38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: MPIO3-0180NMIM
CURRENT APPLICATION NUMBER: US/10/377, 072
CURRENT FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/895, 860
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 370
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/723, 806
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/187, 455
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/843, 297
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199, 801
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: US 09/861, 801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205, 508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/816, 494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/815, 419
PRIOR FILING DATE: 2001-03-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 3544
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (589) ... (2586)
US-10-377-072-25

Alignment Scores:
Pred. No.: 5 06e-201 Length: 3544
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conserved: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
Gaps: 0
DB: 16

US-10-029-345a-109_COPY_1_302 (1-302) x US-10-377-072-25 (1-3544)

QY 1 MetalHhigLumectilegylThrglnleValThrglnrghleuValAlaleuLeuGlu 20
DB 589 ATGGCCCATAGATGATGAACTCAATGTTACTGAGAGGTTGGCGCTCTCGCGAA 648
QY 21 SerGlyThrglnlyrValleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 649 AGTGAAGACGAAAGAGTGTGTAATGATGATGATGATGATGATGATGATGATGATGAT 708
QY 41 HiiLileuLunleuLaleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 709 CACATTTTGAAGCCATTAATATCACTGCTCCAGCTTAATGAAGGCAAGGTTGCAACAG 768
QY 61 AspLysValleuLeuThrglnleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 769 GACAAAGTGTATTAATACAGAGCTCATCAATCCAGCAATACATAGTTGACATTGAT 828
QY 81 CysSerGlnlyrValValValValValValValValValValValValValValVal 100
DB 829 TGCAGTCAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888
QY 101 AspCysPheLeuThrValleuLeuGlyLysLeuGlyLysSerPheLeuValHisLeu 120

DB 889 GACTGTTTTCTCACTGTAATCTTGGGTAACTGAGAAAGCTTCACTGTTCACTG 948
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGlyGlySer 140
DB 949 CTTGAGAGGAGGTTGCTGAGTCTCTGCTGTTTCTCGGCGCTCTGTAGAGAAATCC 1008
QY 141 ThrLeuValProThrCysHisSerGlnProCysLeuProValAlaAniLeuProThr 160
DB 1009 ACTTATGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
QY 161 ArgLysLeuProAsnLeuThrglnleuGlyCysGlnArgAspValleuLeuLeuLeu 180
DB 1069 CGAATTTCTCCCAATCTTATCTTGGCTCCAGCAGATGCTCTCAACAAGAGCTGATG 1128
QY 181 GlnGlnAenglyYlleglyTyrrValleuAlaAsertyrThrCysProIysProAspPhe 200
DB 1129 CAGCAGAAATGGATGGTATGTTATGTTAAATGACAGAAATCTGTCACAAAGCTGACTTT 1188
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlyLysIleLeu 220
DB 1189 ATCCCGAGTCTCAATTTCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
QY 221 ProTyrLeuAspLysSerValAspPheIleGlyLysAlaAlaSerAenglyCysVal 240
DB 1249 CCGTGGTTGACAAATCAGTATGTTCAATTTGATGAAAGCAAAAGCTCCATGATGATGTT 1308
QY 241 LeuValHisCysLeuAlaGlyLysSerArgSerAlaThrIleAlaIleAlaIleMet 260
DB 1309 CTAGTGACCTGTTAGTCTGGAATCTCCCGCTCCGACCATGCTTCCCTCAATCATG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrrArgPheValIleGlyLysArgProThr 280
DB 1369 AAGAGATGACATGCTTATGATGAGCTTACAGATTTGTGAAAGAAAAAGACTTACT 1428
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrrGlyLysIleLeu 300
DB 1429 ATATCTCCAACTTCAATTTTCTGGGCAATCTCTGATGATGAGAAAGATTAAAGAC 1488
QY 301 GlnThr 302
DB 1489 CAGACT 1494

RESULT 9
US-10-425-114-26234
Sequence 26234, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FUS REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425, 114
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 26234
LENGTH: 3625
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB4119-028-H6_FU1
US-10-425-114-26234

Alignment Scores:
Pred. No.: 5.26e-201 Length: 3625
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conserved: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0

DB: 13 Gaps: 0

US-10-029-345a-109_copy_1_302 (1-302) x US-10-425-114-26234 (1-3625)

QY 1 MetAlahisgluMetilegIYThrGlnIleValThrgluArgleuValAlaLeuLeuGlu 20

Db 692 ATGGCCCATGAGATGATGGAATCAAAATGTTACTGAGAGGTTGGTGGCTCTGGAA 751

QY 21 SerGlyThrgluysValleuLeuIleaspSerArgProPheValGluTyrAsnThrSer 40

Db 752 AGTGAACGGAAAAAGTGGCTGTAATGATGACCGGCGCATTTGGAAATCAATAACATCC 811

QY 41 HisIleuGlnAlaIleAsnIleAsnGlySerIleuMetIysArgArgLeuGln 60

Db 812 CACATTTTGAAGCATTAATATCACTGCTCCAGCTTAAGAACGAAAGTTGCAACAG 871

QY 61 AsplysValleuIleThrgluLeuIleGlnHisSerAlaYshIlyValAspIleasp 80

Db 872 GACAAAGTGAATTAATACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 931

QY 81 CysSerGlnIysValIleValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100

Db 932 TGCAGTCAGAAAGTTGATGATTACGATCAAGCTCCAAAGTGTGCTCTCTCTTCA 991

QY 101 AspCysPheLeuThrValleuLeuGlyIlySerGlnIysSerPheAsnSerValHisIleu 120

Db 992 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGACCTTCACTCTTCACTG 1051

QY 121 LeuAlaGlyIlyPheAlaGluPheSerArgCysPheProGlyLeuGlyIlySer 140

Db 1052 CTTCAGAGTGGGTTGGTGAAGTCTCTGTTGTTCCCTGCGCTCTGGAAGAAATCC 1111

QY 141 ThrIleuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160

Db 1112 ACTCTAGTCCCTACCTGATTTCTGAGCTTCTGCTTACCTGTTGCCAACATTGGGCCAAC 1171

QY 161 ArgIleuProAsnLeuIlyrLeuGlyCysGlnArgAspValLeuAsnIlySGluLeuIle 180

Db 1172 CGAATTTCTCCATCTTATCTTGGCTGCCAGGAGATGTCCTCAACAAAGAGCTGATG 1231

QY 181 GlnGlnAsnGlyIlyegIYrValleuAsnAlaSerTyrThrCysProIysProAspPhe 200

Db 1232 CAGAGATGGATGGATTGATGTATTAATCCAGCATATACCTGTCMAAGCTGACTTT 1291

QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIlyIleu 220

Db 1292 ATCCCCGAGTCTCATTTCTGCTGCTGCTGATGATGACACTTTTGGAGAAATTTTG 1351

QY 221 ProTrpLeuAspIysSerValAspPheIleGluIysAlaIysAlaSerAsnGlyCysVal 240

Db 1352 CCGAGTTGGACAATATGATGATTTTCAATGAAAGCAAAAGCTCCAAATGATGTGT 1411

QY 241 LeuValHisCysLeuAlaGlyIlySerArgSerAlaThrIleAlaIleAlaTyrIleMet 260

Db 1412 CTAGTGCATCTGTTAGCTGGGATCTCCGCTCCGCCACATCGCTATCGCTTACATCATG 1471

QY 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValIlySGluIysArgProThr 280

Db 1472 AAGAGATGACATGTCTTTAGATGAACTTACAGATTTTGGAAAGAAAAAGCTTACT 1531

QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIlyrGluIlyIleLysAsn 300

Db 1532 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCCGAGCTATGAGAAAGATTAAGAAC 1591

QY 301 GlnThr 302

Db 1592 CAGACT 1597

RESULT 10

US-10-343-357-17

; Sequence 17, Application US/10343357

; Publication No. US20040058341A1

; GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.Tom

APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi

APPLICANT: YAO, Monique G.; BURROD, Neil

APPLICANT: WANG, Yumei B.; STEWART, Elizabeth A.

APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.

APPLICANT: LEE, Ernestine A.; HARALIA, April J.A.

APPLICANT: LU, Dying Anna M.; TRIBOULEY, Catherine M.

APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Marian R.

APPLICANT: NGUYEN, Henry; WARREN, Bridget A.

APPLICANT: NGUYEN, Daniel B.; CHAWLA, Narinder K.

APPLICANT: KEARNEY, Liam

FILE REFERENCE: PROTEIN PHOSPHATASES

CURRENT APPLICATION NUMBER: US/10/343,357

CURRENT FILING DATE: 2003-01-28

PRIOR APPLICATION NUMBER: PCT/US01/23716

PRIOR FILING DATE: 2001-07-26

PRIOR APPLICATION NUMBER: US 60/221,679

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: US 60/223,272

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: US 60/224,309

PRIOR FILING DATE: 2000-08-10

PRIOR APPLICATION NUMBER: US 60/226,728

PRIOR FILING DATE: 2000-08-18

PRIOR APPLICATION NUMBER: US 60/229,254

PRIOR FILING DATE: 2000-08-30

PRIOR APPLICATION NUMBER: US 60/231,366

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PERL Program

SEQ ID NO 17

LENGTH: 3766

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CBI

US-10-343-357-17

Alignment Scores:

Pred. No.: 5,61e-201 Length: 3766

Score: 1552.00 Matches: 300

Percent Similarity: 99.67% Conservative: 1

Best Local Similarity: 99.34% Mismatches: 1

Query Match: 99.23% Indels: 0

DB: 13 Gaps: 0

US-10-029-345a-109_copy_1_302 (1-302) x US-10-343-357-17 (1-3766)

QY 1 MetAlahisgluMetilegIYThrGlnIleValThrgluArgleuValAlaLeuLeuGlu 20

Db 538 ATGGCCCATGAGATGATGGAATCAAAATGTTACTGAGAGGTTGGTGGCTCTGGAA 597

QY 21 SerGlyThrgluysValleuLeuIleaspSerArgProPheValGluTyrAsnThrSer 40

Db 598 AGTGAACGGAAAAAGTGGCTGTAATGATGACCGGCGCATTTGGAATCAATAACATCC 657

QY 41 HisIleuGlnAlaIleAsnIleAsnGlySerIleuMetIysArgArgLeuGln 60

Db 658 CACATTTTGAAGCATTAATATCACTGCTCCAGCTTAAGAACGAAAGTTGCAACAG 717

QY 61 AsplysValleuIleThrgluLeuIleGlnHisSerAlaYshIlyValAspIleasp 80

Db 718 GACAAAGTGAATTAATACAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 777

QY 81 CysSerGlnIysValIleValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100

Db 778 TGCAGTCAGAAAGTTGATGATTACGATCAAGCTCCCAAGATGTGCTCTCTCTTCA 837

QY 101 AspCysPheLeuThrValleuLeuGlyIlySerGlnIysSerPheAsnSerValHisIleu 120

Db 838 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGACCTTCACTGTTCACTG 897

QY 121 LeuAlaGlyGlyPheAlaGlyPheSerArgCysPheProGlyLeuGlyGlySer 140
| | | | |
Db 898 CTTGAGAGTGGGTTTGGTGGTCTCTCGTTGTTCCCTGCTGTGAAGAAATCC 957
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAniIleGlyProThr 160
| | | | |
Db 958 ACTCTAGTCCCTACCTGATTTCTGAGCCTTGCTTACTGTTGCCAATTTGGGCCAAC 1017
QY 161 ArgIleLeuProAsnLeuTyrIleuGlyCysGlnArgAspValLeuAniIleGlyLeu 180
| | | | |
Db 1018 CGAATCTTCCCAATCTTAATCTTGGCTGACAGATGCTCTCAACAAGAGCTGATG 1077
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
| | | | |
Db 1078 CAGCAAGATGGATTTGGTTATGTGTTAAATGCCAGCAATCTGCTCAAGAGCTGATG 1137
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlyIleLeu 220
| | | | |
Db 1138 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 1197
QY 221 ProTyrLeuAspLysSerValAspPheIleGlyLeuAlaLysAlaSerArgGlyCysVal 240
| | | | |
Db 1198 CCGTGGTTGGACAAATCAGTAGATTTCTTGAAGAAAGCAAAAGCTCCCAATGATGTGT 1257
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
| | | | |
Db 1258 CTAGTGACCTGTTTACTGGATCTCCCGCTCCGCAACCTGCTTATGCTTACATCATG 1317
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGlyLysArgProThr 280
| | | | |
Db 1318 AAGAGATGAGCATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAGAACCTACT 1377
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysIleLysAsn 300
| | | | |
Db 1378 ATATCTCCAAATCTTAATTTCTGGGCGCACTCTGAGACTATGAGAAAGATTAAAGAAC 1437
QY 301 GlnThr 302
| | | | |
Db 1438 CAGACT 1443

RESULT 11
US-10-648-593-115
; Sequence 115, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648, 593
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 115
; LENGTH: 4790
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-648-593-115

Alignment Scores:
Pred. No.: 8,47e-201
Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Query Match: 99.23%
17
Length: 4790
Matches: 300
Conservative: 1
Mismatch: 1
Indels: 0
Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x US-10-648-593-115 (1-4790)

QY 1 MetAlaHisGlyMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
| | | | |
Db 184 ATGGCCCATGAGATGATTTGGAATCAATTTTACTGAGAGGTTGGTGGCTGCTGGAA 243
QY 21 SerGlyThrGlyLysValLeuLeuIleAspSerArgProPheValGlyTyrAsnThrSer 40
| | | | |
Db 244 AGTGAAGACGAAAGAAATGCTGCTTAATTTGATAGCCGCGCAATTTGTGAAATCAATACATCC 303
QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGln 60
| | | | |
Db 304 CACATTTGGAGCCATTAATATCACTCTCCAGCTTATGAAAGGAGGTTGCAACAG 363
QY 61 AspLysValIleuIleThrGlyLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
| | | | |
Db 364 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCCAAACATTAAGTTGATCAT 423
QY 81 CysSerGlnLysValValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
| | | | |
Db 424 TGCAGTCAGAGAGTTTATGATTACGATCAAGCTCCCAAGATTTGCTCTCTCTCA 483
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu 120
| | | | |
Db 484 GACTGTTTCTCAGCTGACTTCTGGGTAACTGGAGAAAGCTTCAACTGTTCACCTG 543
QY 121 LeuAlaGlyGlyPheAlaGlyPheSerArgCysPheProGlyLeuGlyGlySer 140
| | | | |
Db 544 CTTGAGAGTGGGTTTCTGATTTCTCTGTTTCCCTGCTGTGAAGAAATCC 603
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAniIleGlyProThr 160
| | | | |
Db 604 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGCTTACTGTTGCCAATTTGGGCCAAC 663
QY 161 ArgIleLeuProAsnLeuTyrIleuGlyCysGlnArgAspValLeuAniIleGlyLeu 180
| | | | |
Db 664 CGAATCTTCCCAATCTTAATCTTGGCTGCGAGAGATGCTCTCAACAAGAGCTGATG 723
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
| | | | |
Db 724 CAGCAAGATGGATTTGTTATGTGTTAAATGCCAGCAATCTGCTCAAGCTGATCTT 783
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlyIleLeu 220
| | | | |
Db 784 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 843
QY 221 ProTyrLeuAspLysSerValAspPheIleGlyLysValAlaSerArgGlyCysVal 240
| | | | |
Db 844 CCGTGGTTGGACAAATCAGTAGATTTCTTGAAGAAAGCAAAAGCTCCCAATGATGTGT 903
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
| | | | |
Db 904 CTAGTGACCTGTTTACTGGATCTCCCGCTCCGCAACCTGCTTATGCTTACATCATG 963
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGlyLysArgProThr 280
| | | | |
Db 964 AAGAGATGAGCATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAGAACCTACT 1023
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysIleLysAsn 300
| | | | |
Db 1024 ATATCTCCAAATCTTAATTTCTGGGCGCACTCTGAGACTATGAGAAAGATTAAAGAAC 1083
QY 301 GlnThr 302
| | | | |
Db 1084 CAGACT 1089

RESULT 12
US-10-094-749-673
; Sequence 673, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAL, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI

```

APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHITO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 673
LENGTH: 2102
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-673

```

```

Alignment Scores:
Pred. No.: 1,88e-200 Length: 2102
Score: 1545.00 Matches: 299
Percent Similarity: 99.34% Conservative: 1
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 98.79% Indels: 0
Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x US-10-094-749-673 (1-2102)

```

```

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 56 ATGGCCCATGAGATGATTTGGAACCAATTTGACTGAGAGTGGTGGCTCTGCTGAAA 115
QY 21 SerGlyThrGluValLeuLeuLeuIleAspSerArgPropheValGluTyrAsnThrSer 40
DB 116 AGTGAACGGAAGAAAGTCTGCTGTAATGATAGCCGCCCATTTTGAGATACAAATACACC 175
QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleuMetIleValArgArgLeuGln 60
DB 176 CACATTTTGAAGCATTAATATCACTGCTCAAGCTTAATGAGCGAAGGTTGCAACG 235
QY 61 AspIleValLeuIleThrGlnLeuIleGlnHisSerAlaValHisIleValAspIleAsp 80
DB 236 GACAAAGGTTAATTAACAGAGCTCATCAAGCATTCAGGAACATTAAGGTTGACATGAT 295
QY 81 CysSerGlnValValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 296 TGCACTCAGAGGTTGATGATTTACATCAAGACTCCAAAGATGTTGCTCTCTCTTCA 355
QY 101 AspCysPheLeuThrValLeuLeuGluIleValLeuGluValSerPheAsnSerValHisLeu 120
DB 356 GACTGTTTTCACATGATCTCTGCGTAAACTGAGGAAGAGCTTCACCTGTTTCACTG 415
QY 121 LeuAlaGlyValPheAlaGluPheSerArgCysPheProGlyLeuCysGluGluValSer 140
DB 416 CTGGAGGAGGTTGCTGAGAGTCTCTGCTGTTTCCCTGGGCTCTGAGAGGAAATCC 475
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 476 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGCTTACCTGTTGCCAACAATGGGCAACC 535
QY 161 ArgIleLeuProAsnLeuTyrIleGluIleCysGlnArgAspValLeuAsnIleGluLeuIle 180

```

```

DB 536 CTAATCTTCCCAATCTTTATCTTGGCTGCACGAGATGCTCTCAACAAGAGCTGATG 595
QY 181 GlnGlnAspGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProIysProAspPhe 200
DB 596 CAGCAGATGGAGATGGTATGATGTTAATCCAGCAATACCTGTCCAAAGCTGACTTT 655
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleLeu 220
DB 656 ATCCCGAGTCTCATTTCTCGCTGCTGCTGATGACAGACTTTGTTGAGAAATTTTG 715
QY 221 ProTyrLeuAspIleSerValIleAspPheIleGlnIleValAlaValAlaSerAspGlyVal 240
DB 716 CCGTGGTTGACAAATCAAGATTTCTTGAAGAAAGCAAGCTTCATGAGATGTT 775
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
DB 776 CTAGTGACATGTTTAACTGGGATCTCCCGCTCCGCCACATTCGCTCATCATCATG 835
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluValArgProThr 280
DB 836 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTTGAAGAAAAAGCACTACT 895
QY 281 IleSerProAsnPheAsnPheLeuGluIleLeuAspTyrGluValIleValAsn 300
DB 896 ATATCTCCAACCTTCAATTTCTGGGCCAATCTCTGAGCTATGAGAGGATTAAGAAC 955
QY 301 GlnThr 302
DB 956 CAGACT 961

```

```

RESULT 13
US-10-296-115-520
Sequence 520, Application US/10296115
Publication No. US20040053248A1
GENERAL INFORMATION:
APPLICANT: Hyeq Inc
TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
FILE REFERENCE: 784PCT
CURRENT APPLICATION NUMBER: US/10/296,115
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: US09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 1478
SEQ ID NO 520
LENGTH: 2966
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(2966)
OTHER INFORMATION: n = a,t,c or g
US-10-296-115-520

```

```

Alignment Scores:
Pred. No.: 4.66e-200 Length: 2966
Score: 1544.00 Matches: 299
Percent Similarity: 99.34% Conservative: 1
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 98.72% Indels: 0
Gaps: 0
DB: 13

```

US-10-029-345A-109_COPY_1_302 (1-302) x US-10-296-115-520 (1-2966)

```

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 23 ATGGCCCATGAGATGATTTGGAACCAATTTGACTGAGAGTGGTGGCTCTGCTGAAA 82
QY 21 SerGlyThrGluValLeuLeuLeuIleAspSerArgPropheValGluTyrAsnThrSer 40
DB 83 AGTGAACGGAAGAAAGTCTGCTGTAATGATAGCCGCCCATTTTGAGATACAAATACATACC 142

```


Qy	41	HisGILeLengUuAlaIleAsnIleAsnCySeSerIySLeuMetLySAAGxIeUgIn	60
Db	143	CACATTTTGGAAAGCCATTATATATCAACTGCTCAAGCTTATGAAAGCAAGTTTCAACAG	2020
Qy	61	AspLySeValIleuIleThrGluIleuIleGlnIleSerAlaIySHIeLySeValAspIleAsp	80
Db	203	GACAAAGTGTATATTCACAGAGCTCAATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT	2626
Qy	81	CysSeSerGlnIySeValIyTyrArgGInSeSerGlnAspValAlaSeLeuSeSer	1000
Db	263	TGAGGTCAGAAAGGTGTAGTTTACATCAATCAAAAGCTCCAAAGATGTTGCTCTCTCTTCA	3222
Qy	101	AspCySePheIeuThrValIleuIleuGlyIySLeuGlnIySeSerPheAsnSeValIleu	1200
Db	323	GACTGTTTTCTCACTACTACTCTGGGTAAACCTGGAAAGAGACTTCACTGTTCACTG	3822
Qy	121	LeuAlaGlyIyPheAlaGlnPheSeSerArgPheProGlyIeUeCySeGlnIyLySeSer	1400
Db	383	CTTGCAAGTGGGTTTCTGAGTTCTCTCGATTCTTCCTGAGCTCTGGAAGAAATTC	4422
Qy	141	ThrIeuValProThrCySeValIeSerGlnProCySLeuProValAlaAsnIleGlyProThr	1600
Db	443	ACCTATGATCCCTTACCTGCAATTCCTCAAGCTTGCTTACCTGTTGCCAAATATGGGCCAAC	5020
Qy	161	ArgIleIeuProAsnIeuTyrIleuGlyCySeGlnArgAspValIleAsnIySeGlnIle	1800
Db	503	CGAATTCCTCCCAATCTTTATCTTGCTGCTGCCAGCAAGATGTCCTCAACAGAGACTATG	5622
Qy	181	GlnGlnAsnGlyIleGlyTyrValIleuAsnAlaSerTyrThrCySeProLyProAspPhe	2000
Db	563	CACACAGATGGAGTGGTATGATGTAAATGCCACATATACCTGHCCAAAGCTGACTTT	6222
Qy	201	IleProGlnIySeSerIyPheIeuArgValProValaAsnAspSerPheCySeGlnIyIleu	2200
Db	623	ATCCCGAGCTCATATTCCTGGGTGGCTGGAATGACAGCTTTTGAAAGAAATTTTG	6822
Qy	221	ProTrpIeuAspIySeSerValAspPheIleGlnIyAlaIySAlaIySeAsnGlyCySeVal	2400
Db	683	CCGGTGTGGACAAATCAAGTATTTCAATGAAAGGAAAGCAAGCTCCAAAGATGTGTT	7422
Qy	241	LeuValHisCySeIeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet	2600
Db	743	CTAGTGACAGTTTAACTGCGGATCTCCGCCCTCGGCACACATCGCTATGCTTACATATG	8020
Qy	261	LyArgIeuAspIeSeSerIeuAspGlnAlaTyrArgPheValIySeGlnIyArgProThr	2800
Db	803	AAAGAGATGGACATGCTTAAATGGAAGCTTACAAATTTGGAAAGGAAAGGCACTACT	8622
Qy	281	IleSeSerProAsnPheAsnPheIeuGlnIyGlnIleuAspTyrGlnIyIySAlaIyAsn	3000
Db	863	ATATCTCCAAACTTCAATTTTCTGGGCCCACTCCCTGACCTATGAGAGAAATTAAGAAC	9222
Qy	301	GlnThr 302	
Db	923	CAGACT 928	
RESULT 14			
US-10-072-012-255			
Sequence 235, Application US/10072012			
Publication No. US20040033493A1			
GENERAL INFORMATION:			
APPLICANT: Tchenev, Vellizar			
APPLICANT: Spytek, Kimberly			
APPLICANT: Zernusen, Bryan			
APPLICANT: Raturajan, Meera			
APPLICANT: Shimkets, Richard			
APPLICANT: Li, Li			
APPLICANT: Gangoli, Beha			
APPLICANT: Padigaru, Muralidhara			
APPLICANT: Anderson, David W.			
APPLICANT: Rastelli, Luca			
APPLICANT: Miller, Charles E.			

```

APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsbrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 225
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
US-10-072-012-255

Alignment Scores:
Pred. No.: 7,28e-195 Length: 2200
Score: 1504.50 Matches: 300
Percent Similarity: 94.95% Conservative: 1
Best Local Similarity: 94.64% Mismatches: 1
Query Match: 96.20% Indels: 15
DB: 13 Gaps: 4

US-10-029-345A-109_COPY_1_302 (1-302) x US-10-072-012-255 (1-2200)
Dh 1 MetAlHisGluMetIleGlyThrGlnIleValI ThrGluArgLeuValAlaLeuLeuGln 20
61 ATGGCCCATGATGATGATGATGAACTCAATTTGTTACTGAGAGGTTGGTCTGCTGGAA 120
Qy 21 SerGlyThrGlnIlyValLeuLeuIleAspSerArgProPheValGlnTyrAsnThrSer 40
Dh 121 AGTGGACCGGAAAAAGTCGTGAATGTGATGACCGGCAATTGGAGATTACAAATCATCC 180
Qy 41 HisIleLeuGlnValIleAsnIleAsnCysSerIlySleMetIlyArgArgLeuGlnGln 60
Dh 181 CACATTTTGGAAACCTTAATATCAACTCTCCAAAGCTTTATGAGCGGAAGGTGCAACG 240
Qy 61 AspIlyValLeuIleThrGlnIleuIleGlnHisSerAlaIlyPheIlyS----- 76
Dh 241 GACAAAGTGTATATTCAGAGCTCATCCGCAATTCAGCGAAACATTAAGTAAACGCTCAG 300
Qy 77 ValAspIleAspCysSerGlnIlySValValValIlyTyrAspGlnSerSerGlnAspValAla 96

```

```

Db      301 GTTGACATGATTGCAGTGCAGAAAGTTGATTACGATCAAGACTCCCAAGATGTTGCC 360
Qy      97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsn 116
Db      361 TCTCTCTCTCCAGACTGTTTTCTCAGTGTATCTTCGGGTAAACGAGGAAAGCTTCAAC 420
Qy      117 SerValHisLeuLeuAla-----GlyGlyPheAlaGluPheSerArgCysPhe 132
Db      421 TCTGTCACTGCTTGACAGGTTTATCTTAGGTGGGTTTCTGAGTTCTCTCGTTGTTTC 480
Qy      133 ProGlyLeuCysGluGlyLysSerThrLeuValProThrCysIleSerGlnProCysLeu 152
Db      481 CTTGGCCCTGTGTGAAGAAATTCACCTTAGTCCCTACCTGCATTTCTCAGCCTTGCTTA 540
Qy      153 ProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrLeuGlyCysGluArg 172
Db      541 CCTGTGGCAACATTTGGGCCCAACCCGAATCTTCCCATCTTATCTTGGCTGCAGAGA 600
Qy      173 AspValLeuAsnLysGlyLeuIleGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSer 192
Db      601 GATGTCCTCAACAAAGAGACTGATGACGACAAATGGAGTTGGTTATGTGTTAAATCCAGC 660
Qy      193 TyrThrCysProLysProAspPheIleProGlnSerHisPheLeuArgValProValAsn 212
Db      661 AATACCTGTCCAAAGCCCTGACTTATCCCGAGTCTCATTTCTGCGGTGCTGTGAT 720
Qy      213 AspSerPheCysGlyLysIleLeuProThrLeuAspLysSerValAspPheIle----- 230
Db      721 GACAGCTTTGTGTGAAAAATTTGGCGGTGGAACAAATCAGATGATTCAATGGTAAAG 780
Qy      231 -----GlyLysAlaLysAlaSerAsnGlyCysValLeuValHisCysLeuAla 246
Db      781 TTGACTTATACAGAGAAAGCAAAAGCTCCCAATGATGTGTTAGTGCACATGTTAGCT 840
Qy      247 GlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLysArgMetAspMetSer 266
Db      841 GGAGTCTCCCGCTCGGCACCATCGCTATCGCTCATCATGAAGAGATGACATGTCT 900
Qy      267 LeuAspGluAlaTyr--ArgPheValLysGlyLysArgProThrIleSerProAsnPhe 285
Db      901 TTAGTGTAGAGCTTACAGAGATTGTGTGAAGAAAAAACCTACTATATCTCCAAACTTC 960
Qy      286 AsnPheLeuGlyGlnLeuLeuAspTyrGlyLysLysIleLysAsnGlnThr 302
Db      961 AATTTTCTGGCCCAACTCTTGACATATGAGAAAGATTAGAACACAGACT 1011

RESULT 15
US-09-964-277-20
; Sequence 20, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucite, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-964-277-20

Alignment Scores:
Pred. No.: 4,75e-152 Length: 3332
Score: 1194.50 Matches: 244
Percent Similarity: 81.19% Conservative: 2
Best Local Similarity: 80.53% Mismatches: 1
Query Match: 76.37% Indels: 56
DB: 9 Gaps: 1

```

```

US-10-029-345a-109_copy_1_302 (1-302) x US-09-964-277-20 (1-3332)
Qy      1 MetAlaHisGlyMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db      562 ATGGCCATGATGATGATTGGAACCTCAATGTTTACTGAGAGGTTGGTGCTGCTGGA 621
Qy      21 SerGlyThrGlyLysValLeuLeuIleAspSerArgProPheValGlyTyrAsnThrSer 40
Db      622 AGTGAACGGAATAAGTGCTGTAATTGATAGCCGGCCATTTGTGAAATCAATACATCC 681
Qy      41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgGlyLeuGln 60
Db      682 CACATTTTGAACCATTAATATCACTGCTCAAGCTTATGAAACCAAGGTTGCAACAG 741
Qy      61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db      742 GACAAAGTTTAATTACAGAGCTCATCCAGCATTTACGCGAAACATAGTTGACATTGAT 801
Qy      81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db      802 TGCAGTCAAGAGTTGATGATTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 861
Qy      101 AspCysPheLeuThrValIleLeuLeuGlyLysSerPheAsnSerValHisLeu 120
Db      862 GACTGTTTCTCACTGTACTTCTGGTAACTGGAGAAAGCTTCAACTGTCTTCACTTG 921
Qy      121 -LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyLysSe 140
Db      922 CTTC----- 925
Qy      140 rThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProth 160
Db      925 ----- 925
Qy      160 rArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
Db      926 -----CAGAGCTGAT 936
Qy      180 eGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAsp 200
Db      937 GCAGCAGAAATGGATTTGTTATGTATGTTAAATGCCAGAAATACCTGTCAAAGCCTGACTT 996
Qy      200 eIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlyLysIle 220
Db      997 TATCCCGAGTCTCATTTCTCGCTGTGCTGTGAATGACAGCTTTGTGAAAAATTTT 1056
Qy      220 uProThrLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCysVa 240
Db      1057 GCCGTGTGTGACAAATCAAGATTCATTTGAGAAAGCAAAAGCTCCCAATGATGATGTCT 1116
Qy      240 IleValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMe 260
Db      1117 TCTAGTGACAGTTTACTGTGGATCTCCCGCTCGCCACCATCTGCTATCGCTCATCAT 1176
Qy      260 tLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGlyLysArgProth 280
Db      1177 GAAGAGATGACATGCTTTAGATGAAGCTTACAGATTGTGTGAAGAAAAAACCACTAC 1236
Qy      280 rIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysLysIleLysAs 300
Db      1237 TATATCTCCAAACTTCAATTTTCTGGCCCAACTCTTGACATATGAGAAAGATTAGAA 1296
Qy      300 nGlnThr 302
Db      1297 CCAGACT 1303

```

Search completed: June 22, 2004, 03:06:44
Job time : 381.269 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 21:10:41 ; Search time 59.9628 seconds
 (without alignments)
 2794.987 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
 Perfect score: 1564
 Sequence: 1 MAHEMIGTIVTERVALLE.....PNNFUGOLDYBKIKKNGT 302

Scoring table:
 BLOSUM62
 Xgapop 10.0 , Xgapext 0.5
 Ygapop 10.0 , Ygapext 0.5
 Fgapop 6.0 , Fgapext 7.0
 Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
 Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
 Maximum DB seq length: 200000000
 Post-processing: Minimum Match 0%
 Listing first 45 summaries

Command line parameters:
 -MODE=frame+ p2n.model -DEV=xlh
 -O=/cgn2.1/USPTO.spool/US10029345/rnatc 21062004.164146.29052/app_query.fasta_1.1294
 -DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPLC=0
 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
 -USER=US10029345 @CNC 1.1 76 @rnatc 21062004.164146.29052 -NCPU=6 -ICPU=3
 -NO MMP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
 -DEF TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
 1: /cgn2.6/prodata/2/ina/5A.COMB.seq.*
 2: /cgn2.6/prodata/2/ina/5B.COMB.seq.*
 3: /cgn2.6/prodata/2/ina/6A.COMB.seq.*
 4: /cgn2.6/prodata/2/ina/6B.COMB.seq.*
 5: /cgn2.6/prodata/2/ina/PCTUS.COMB.seq.*
 6: /cgn2.6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1552	99.2	1998	4	US-09-816-494-3	Sequence 3, App1
2	1552	99.2	3544	4	US-09-816-494-1	Sequence 1, App1
3	988	63.2	2377	4	US-09-920-668-3	Sequence 3, App1
4	469	30.0	1830	4	US-09-557-921-1	Sequence 1, App1
5	467.5	29.9	2109	4	US-09-016-434-1135	Sequence 1135, App
6	467.5	29.9	2109	4	US-09-023-655-946	Sequence 946, App
7	426	27.2	2240	4	US-09-016-434-1100	Sequence 1100, App
8	425	27.2	944	4	US-09-371-671B-10	Sequence 10, App1
9	425	27.2	1238	2	US-08-530-290-11	Sequence 11, App1
10	425	27.2	1238	4	US-09-702-705-803	Sequence 803, App
11	425	27.2	1238	4	US-09-736-457-803	Sequence 803, App
12	425	27.2	1238	4	US-09-614-124B-803	Sequence 803, App

13	425	27.2	1238	4	US-09-671-325-803	Sequence 803, App
14	425	27.2	1238	4	US-09-589-184-803	Sequence 803, App
15	425	27.2	2064	4	US-09-702-705-825	Sequence 825, App
16	425	27.2	2064	4	US-09-736-457-825	Sequence 825, App
17	425	27.2	2064	4	US-09-614-124B-825	Sequence 825, App
18	425	27.2	2064	4	US-09-671-325-825	Sequence 825, App
19	425	27.2	2064	4	US-09-589-184-825	Sequence 825, App
20	425	27.2	2109	4	US-09-702-705-826	Sequence 826, App
21	425	27.2	2109	4	US-09-736-457-826	Sequence 826, App
22	425	27.2	2109	4	US-09-614-124B-826	Sequence 826, App
23	425	27.2	2109	4	US-09-671-325-826	Sequence 826, App
24	425	27.2	2109	4	US-09-589-184-826	Sequence 826, App
25	423.5	27.1	1993	2	US-08-990-379-2	Sequence 2, App1
26	415	26.5	2000	4	US-09-016-434-1291	Sequence 1291, App
27	408	26.1	2303	4	US-09-922-146-3	Sequence 3, App1
28	382	24.4	1987	2	US-08-990-379-1	Sequence 1, App1
29	380	24.3	279	4	US-09-016-434-91	Sequence 91, App1
30	357	22.8	1619	4	US-09-702-705-801	Sequence 801, App
31	357	22.8	1619	4	US-09-736-457-801	Sequence 801, App
32	357	22.8	1619	4	US-09-614-124B-801	Sequence 801, App
33	357	22.8	1619	4	US-09-671-325-801	Sequence 801, App
34	357	22.8	1619	4	US-09-589-184-801	Sequence 801, App
35	357	22.8	1619	4	US-09-702-705-804	Sequence 804, App
36	357	22.8	1619	4	US-09-736-457-804	Sequence 804, App
37	357	22.8	1619	4	US-09-614-124B-804	Sequence 804, App
38	357	22.8	1619	4	US-09-671-325-804	Sequence 804, App
39	357	22.8	1619	4	US-09-589-184-804	Sequence 804, App
40	295.5	18.9	539	4	US-09-389-681-311	Sequence 311, App
41	295.5	18.9	539	4	US-09-620-405B-311	Sequence 311, App
42	295.5	18.9	539	4	US-09-339-338-311	Sequence 311, App
43	295.5	18.9	539	4	US-09-433-826B-311	Sequence 311, App
44	295.5	18.9	539	4	US-09-604-287A-311	Sequence 311, App
45	295.5	18.9	539	4	US-09-834-759-311	Sequence 311, App

ALIGNMENTS

RESULT 1
 US-09-816-494-3
 ; Sequence 3, Application US/09816494
 ; Patent No. 6664089
 ; GENERAL INFORMATION:
 ; APPLICANT: Meyers, Rachel A.
 ; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: 10448-030002
 ; CURRENT APPLICATION NUMBER: US/09/816,494
 ; CURRENT FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: US 60/191,858
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1998
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-816-494-3

Alignment Scores:
 Pred. No.: 8.04e-195
 Score: 1552.00
 Percent Similarity: 99.67%
 Best Local Similarity: 99.34%
 Query Match: 99.23%
 DB: 4
 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-816-494-3 (1-1998)

Qy	1	MetAlHIGLumetIleGlyThrGlnIleValThrGluArgValAlaLeuGlu	20
Db	1	ATGGCCCATGAGTGGATGGAACTCAAAATGTACTAGAGGTTGGTGGCTGCGGAA	60
Qy	21	SerGlyThrGluValIleuLeuIleAspSerArgProPheValGluIyrAenThrSer	40

```
Db 61 AGTGAACGAAAAAGCTCTTAATTGATGCGGCGCATTTGGGAATCAATACATCC 120
QY 41 H1S1LEUAGLUALI1LEASN1EASN1CYSESLYSEUWELYSARGA1GLEU1GIN 60
Db 121 CACATTTGGAAAGCCATTAATATCACTGCTCCAAAGCTTATGAAGCAAGTTGCAACG 180
QY 61 ASPLYSVALLEU1LETHRG1LEU1LEGLN1HSER1ALYS1SLYSVALASP1LEAP 80
Db 181 GACAAAGCTTATTAATCAAGAGCTATCCAGCATTCACGAAACATTAAGTTGACATTTGAT 240
QY 81 CYSSEGLNYSVALVAL1VAL1TYRASPGLNSERSEGLN1ASPVAL1ASERLEUSER 100
Db 241 TGCAGTCAGAAAGTTGATGATTACGATCAAAAGCTCCAAATGTTCCCTCTCTCTTCA 300
QY 101 ASPCYSPHELEUTHRVALLEU1LEU1LEU1LYS1LEU1LYS1SERPHEASERVALH1SLEU 120
Db 301 GACTGTTTCTCACTGACTTCTGGGTAAACTGGAGAAAGCTTCAACTCTGTTCACTG 360
QY 121 LEUVALAGLYPHEALAGUPHESERARGCYSPHEPROGLYLEUCYSG1UGLYYSER 140
Db 361 CTGACAGTGGGTTTGCTGAGTTCTCTCGTTGTTTCCCTGGCTCTGTGAGAAATCC 420
QY 141 THRLEUVALPROTHR1CYSL1ESERGLNPROCYSELEUPROVAL1AASN1LEGLYPROTHR 160
Db 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCCTTGCTACTGTTGCCAATGGGCCAAC 480
QY 161 ARG1LEUPROASN1LEUTHR1LEU1LYS1ARG1ARG1ASPVAL1LEU1LEU1LE 180
Db 481 CGAATCTTCCCAATCTTTATCTTCTGCTGCGCGAGATGCTCTCAACAGAGCTGATG 540
QY 181 G1NG1N1ASNG1Y1LEGLYTYRVALLEU1AASN1ASERTYR1THR1CYSPROLYSPROASPPHE 200
Db 541 CAGCAAAATGGGATTTGATGTTAATGCCAGCAATACCTGTCAAAGCTTGACTTT 600
QY 201 ILEPROGL1SER1SPHELEUARGVAL1PROVAL1AASN1SERPHECYSG1ULYS1LEU 220
Db 601 ATCCCGAATCTCATTTCTCGCGTGGCTGTGAATGACAGCTTTTGAGAGAAATTTTG 660
QY 221 PROTPLEUASP1YSERSEVAL1ASPPHE1LEGLYLSA1ALYS1ASER1CYSPHE 240
Db 661 CCGTGGTTGCAAAATCACTAGATTTCAATTGAGAAAGCAAAAGCTCCAAAGATGATG 720
QY 241 LEUVALH1SCYSEUVALAGLY1LESEARGSER1ATHR1LEA1LEA1TYR1LEU 260
Db 721 CTAGTGCACTGTTTGGTGGATCTCCGCTCCGCGCACCATGCTATGCTTACATCATG 780
QY 261 LYSARGMETASPMETSERLEUASPGLUAL1ATYRARGPHEVAL1YSGL1ULYARGPROTHR 280
Db 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTGTTGAAAGAAAGAAAGCACTACT 840
QY 281 ILESERPROASN1PHEASN1PHEU1LYS1LEU1LEU1ASP1YRGLY1SLYS1LEYSASN 300
Db 841 ATATCTCCAAATCTCAATTTCTGGGCCAAGCTCTGAGACTATGAGAAAGAAATTAAGAAC 900
QY 301 G1N1THR 302
Db 901 CAGACT 906
```

RESULT 2

```
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 3862 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2583)
US-09-816-494-1

Alignment Scores:
Pred. No.: 2,1e-194 Length: 3544
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 4 Gaps: 0

US-10-029-345a-109_copy_1_302 (1-302) x US-09-816-494-1 (1-3544)
QY 1 MetAlaH1eGLU1eT1eGL1YThR1G1N1LeVAlThR1G1UArgPHeVAl1A1eU1eU1G1U 20
Db 589 ATGGCCCATGAGATGATGATGGAATGGAATCAAAATTTGTTACTGAGAGTTGGCTGCTGGA 648
QY 21 SERGLYTHRGLYLSYVALLEU1LEU1LEU1LEU1SERARGPROPHEVAL1GLUTYRASN1THR 40
Db 649 AGTGAACGAAAAAGGCTGCTTAATTTGATGACCGGCCCATTTGGGAATACAAATACATCC 708
QY 41 H1S1LEUAGLUALI1LEASN1LEASN1CYSESLYSEUWELYSARGA1GLEU1GIN 60
Db 709 CACATTTGGAAAGCCATTAATATCACTGCTCCAAAGCTTATGAAGCAAGTTGCAACG 768
QY 61 ASPLYSVALLEU1LETHRG1LEU1LEGLN1HSER1ALYS1SLYSVALASP1LEAP 80
Db 769 GACAAAGTTTATTAATCAAGACTCAACGATTCAGCAATTCAGCAAAATAGGTTGACATTTGAT 828
QY 81 CYSSEGLNYSVALVAL1VAL1TYRASPGLNSERSEGLN1ASPVAL1ASERLEUSER 100
Db 829 TGCAGTCAGAAAGTTGATGATTACATCAAAAGCTCCAAAGATGTTCCCTCTCTCTTCA 888
QY 101 ASPCYSPHELEUTHRVALLEU1LYS1LEU1LYS1SERPHEASN1VALH1SLEU 120
Db 889 GACTGTTTCTCACTGACTTCTGCGTAACTCGGAAAGAGTTCACTGTTCACTG 948
QY 121 LEUVALAGLYPHEALAGUPHESERARGCYSPHEPROGLYLEUCYSG1UGLYYSER 140
Db 949 CTGAGAGGTGGTGGTGGATTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 1008
QY 141 THRLEUVALPROTHR1CYSL1ESERGLNPROCYSELEUPROVAL1AASN1LEGLYPROTHR 160
Db 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGTACTGTTGCCAATTTGGGCCAAC 1068
QY 161 ARG1LEUPROASN1PHEASN1PHEU1LYS1LEU1LYS1SERPHEASN1VALH1SLEU 180
Db 1069 CGAATCTTCCCAATCTTATCTTGGCTGCGAGAGATGTCCTCAACAGAGACTGATG 1128
QY 181 G1NG1N1ASNG1Y1LEGLYTYRVALLEU1AASN1ASERTYR1THR1CYSPROLYSPROASPPHE 200
Db 1129 CAGCAGATGGGATTTGATGTTGTTAATGCCAGCAATACCTGTCCAAAGCCGATCTTT 1188
QY 201 ILEPROGL1SER1SPHELEUARGVAL1PROVAL1AASN1SERPHECYSG1ULYS1LEU 220
Db 1189 ATCCCGAATCTCATTTCTCGCTGCTGCTGATGATACAGCTTTTGTGAGAAATTTTG 1248
QY 221 PROTPLEUASP1YSERSEVAL1ASPPHE1LEGLYLSA1ALYS1ASERANG1CYSEVAL 240
Db 1249 CCGTGGTTGCAAAATCAAGATTTATTAAGAAACCAAAAGCTCCCAATGATGATGTT 1308
QY 241 LEUVALH1SCYSEUVALAGLY1LESEARGSER1ATHR1LEA1LEA1TYR1LEU 260
Db 1309 CTAGTGCACTGTTTAGCTGGGATTCCTCCGCTCCGCAACATTCCTATGCTTACATCATG 1368
QY 261 LYSARGMETASPMETSERLEUASPGLUAL1ATYRARGPHEVAL1YSGL1ULYARGPROTHR 280
```

```
Db 1369 AAGAGATGACATCTCTTTAGATGAGCTTACAGATTCTTGAAAGAAAAGACCTACT 1428
Qy 281 IISerProAnPheAnPheLeuGlyGlnLeuLeuAspTyrGluYslyIleYsAn 300
Db 1429 ATATCTCAAACTTCAATTTCTGGGCAACTCTGAGCATATGAGAAGATTAAAGAAC 1488
Qy 301 GlnThr 302
Db 1489 CAGACT 1494

RESULT 3
US-09-920-668-3
; Sequence 3, Application US/09920668
; Patent No. 6482644
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowbert
; APPLICANT: Bret P. Monla
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 8 EXPRESSION
; FILE REFERENCE: RTS-0246
; CURRENT APPLICATION NUMBER: US/09/920,668
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)...(2012)
US-09-920-668-3

Alignment Scores:
Pred. No.: 4,67e-120 Length: 2377
Score: 988.00 Matches: 190
Percent Similarity: 81.40% Conservative: 55
Best Local Similarity: 63.12% Mismatches: 54
Query Match: 63.17% Indels: 2
DB: Gaps: 2

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-920-668-3 (1-2377)
Qy 1 MetAlaHisGluMetIleGlyThrGlnIleVal---ThrGlnArgLeuValAlaLeuLeu 19
Db 135 ATGGCTGGGAGCGGCTCCGAGAGAGTATGATGCCAAGAGCTGGCCAGCTGCTG 194
Qy 20 GluSerGlyThrGluYsValLeuLeuIleAspSerArgProPheValGluYrAsnThr 39
Db 195 CGGGGCGGCGCTGGGGGCGGCTGGTGCATGACAGCGGCTCTTCGTGGAGTACACAGC 254
Qy 40 SerHisIleLeuGlnAlaIleAsnIleAsnCysSerIleLeuMetIleYsArgTgLeuGln 59
Db 255 TGGCATGTGCTCAGCTCCGTCACATCTGCTCCAGCTGGAGTGAAGCGGCGCTGCAG 314
Qy 60 GlnAspIleYsValLeuIleThrGluLeuIleGlnHisSerIleAlaYsHisValAspIle 79
Db 315 CAGGGCAGAGGTGACATTTGGGAGGCTCATCAGCCGCTGCACGCGCTGAGGGCT 374
Qy 80 AspCysSerGlnIleYsValIleValIleYrAspGlnSerSerGlnAspValAlaSerLeuSer 99
Db 375 AGGAGACCAAGACAGAGCTGGTGTCTATGACCAAGACAGCGGAGAGCGGACGCTGGCC 434
Qy 100 SerAspCysPheLeuThrValLeuLeuGlyIleYsLeuGlnIleYsSerPheAsnSerValHis 119
Db 435 GCAAGACAGCTTCTCTCATCTGCTGCTGAGCAAGCTGAGCGGCTTCCAGAGGTGGCC 494
Qy 120 LeuLeuAlaIleYsIlePheAlaGlnPheSerArgCysPheProGlyLeuYsGluGlyYs 139
Db 495 ATCTCACTGGGGGCTTGGCCACTTCTCTCTCTCCCGGCTTCCCGGCTTCCGAGGGGAG 554
Qy 140 ---SerThrIleValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGly 158
Db 555 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 614
```

```
Qy 159 ProThrArgIleLeuProAnPheAnPheLeuGlyCysGlnArgAspValLeuAsnYsGlu 178
Db 615 CTGACCCGATCTGCTCCTCACTTACCTGGGCTCGAGAGAGACGCTTAAACAGAT 674
Qy 179 LeuIleGlnIleAsnGlyIleGlyYrValIleAsnAlaSerYrThrCysProYsPro 198
Db 675 CTGATAGCCCAAAATGGATTAGCTACGCTTCAAGCCAGCAACTCTGCCCAAGCCT 734
Qy 199 AspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluYs 218
Db 735 GACTTCATCTGCGAGAGCCGCTTCATGCGGCTCCCATCAACACAACTACTCTGAAAA 794
Qy 219 IleLeuProThrLeuAspIleYsSerValAspPheIleGlnIleYsAlaIleValAlaSer 238
Db 795 CTGCTCCCTCGGTGAGCAAGTTCATCGAGTTCATGATTAAACCAAGCTCTCCAGTGC 854
Qy 239 CysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaYr 258
Db 855 CAAGTATGCTCATCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 914
Qy 259 IleMetIleArgMetAspMetSerLeuAspGlnAlaIleYrArgPheValIleYsGluYsArg 278
Db 915 ATCATGAGACCATGAGGATGCTCTCCGAGAGAGCTTACAGGTTCCGTGAAGAGACAGCGC 974
Qy 279 ProThrIleSerProAnPheAnPheAnPheLeuGlyGlnLeuLeuAspTyrGluYsIle 298
Db 975 CCGTCCATCTCGGCCCACTTCACTTCTGGGCGAGCTGCTGGAGTACGAGCGCAGCTG 1034
Qy 299 Yrs 299
Db 1035 AAG 1037

RESULT 4
US-09-557-921-1
; Sequence 1, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Wei, Bo
; APPLICANT: Lucie, Ralf M.
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125,416
; CURRENT APPLICATION NUMBER: US/09/557,921
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-557-921-1

Alignment Scores:
Pred. No.: 1,45e-51 Length: 1830
Score: 469.00 Matches: 107
Percent Similarity: 57.72% Conservative: 65
Best Local Similarity: 35.91% Mismatches: 96
Query Match: 29.99% Indels: 30
DB: Gaps: 7

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-557-921-1 (1-1830)
Qy 27 LeuLeuIleAspSerArgProPheValGluYrAsnThrSerHisIleLeuGlnAlaIle 46
Db 603 GTCATCATTTAGCTGAGGCCCTTCATGAGTACAAACAGTCAATCCAAAGAGCTGTC 662
Qy 47 AsnIleAsnCysSer---IleLeuMetIleYsArgTgLeuGlnIleAspIleValIle 65
Db 663 CACATTAACTGTGGCATTAAGTCAACCGCGGAGACTCAGAGAGGCAAGTCACTGCT 722
Qy 66 ThrGluLeuIleGln---HisSerIleAlaYsHisValAspIleAspCysSerGlnYs 84
Db 723 CTAGACTTATTTCTGTAGAGAGCAAGACTCTTTCAAGAGATCTTTTCCAAAGAA 782
```

```
Qy 85 ValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeu 104
Db 783 ATTATAGTTTATGATGAAATATACCAATGACCAAGCCAGTGAAGCCCTCCACCACTT 842
Qy 105 ThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeuLeuAlaGly 124
Db 843 CACATAGCTCTGAGTCCCTGAAAGAGAAAGCAAGAACTCTGCTGTGAAAGGTGA 902
Qy 125 PheAlaGlnPheSerArgCysPheProGlyLeuGlyGlyLys----- 139
Db 903 CTTAGTAGTTTAAACAGAACCAATGAAAACTCTGTGCAACTCCCTCCAGCTCCAAAG 962
Qy 140 -----SerThrLeuValProThrCysIleSer 148
Db 963 TGCAGGAGTGGGGGGGGGGCATCCGGGCTGAGCTTGCATCCTCAACCCATC--- 1019
Qy 149 GlnProCysLeuPro---ValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyr 167
Db 1020 ---CCACACACCTGATGAGAGACGCTGAGCTCACCCCATCTTGCCCTCTGCTC 1076
Qy 168 LeuGlyCysGlnAlaAspValLeuAsnLysGlnLeuIleGlnAsnGlyIleGlyTyr 187
Db 1077 CTGGCAATGAGCAGATGCTCAGACCTGAGACCAATGACGCGCTGAACATCGGCTAC 1136
Qy 188 ValLeuAsnAlaSer-----TyrThrCysProLysProAspPheIlePro 202
Db 1137 GTCATCAACGTCACACATCATCTTCCCTCTACCACTATGAAAGGCTCTTC----- 1190
Qy 203 GluSerHisPheLeuArgValProValAsnAspSerPheCysGlyLysIleLeuProTyr 222
Db 1191 ---AACTACAAGCGGCTGCCAGCCACTGACAGCAACAAGCAACTCGCGCAGTAC 1244
Qy 223 LeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCysValLeuVal 242
Db 1245 TTTAAGAGGCTTTTGAAGTATTGAGAGAGCTCAACAGTGTGGAAGGGCTTCTCATC 1304
Qy 243 HisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIleAlaIle 262
Db 1305 CACGCGCAGGCTGGGGTCTCCCGCTCCGCCACCATCGTCATCGCTTACTGATGAAGAC 1364
Qy 263 MetAspMetSerLeuAspGlnAlaIleTyrArgPheValLysGlyLysArgProThrIleSer 282
Db 1365 ACTCGATGACCACTGACGATGCTTAATTTGTCAAAGGCAAGCAACCAATATATCTCC 1424
Qy 283 ProAsnPheAsnPheLeuGlyGlnLeuAsnAspTyrGlyLysIleLysAsn 300
Db 1425 CCAAACTTAACCTTATGAGGAGGAGTGTCTAGAGTTGAGGAAGACCTAAACAC 1478

RESULT 5
US-09-016-434-1135
; Sequence 1135, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
```

```

; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2109 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1418933
; US-09-016-434-1135

Alignment Scores:
Pred. No.: 2,91e-51 Length: 2109
Score: 467.50 Matches: 110
Percent Similarity: 53.56% Conservative: 63
Best Local Similarity: 34.06% Mismatches: 105
Query Match: 29.89% Indels: 45
DB: Gaps: 7

US-10-029-345a-109_copy_1_302 (1-302) x US-09-016-434-1135 (1-2109)
Qy 19 LeuGluSerGlyThrGlyLysValLeuIleAspSerArgProPheValGlyTyrAsn 38
Db 430 CTGAGCTGGGCAACAGACGCGCTGCTGATGACTGCGCGGACGAGATATACGAG 489
Qy 39 ThrSerHisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgTyrLeu 58
Db 490 TCGTCGACATCCAGTCGCGCATCAACGTCGCCATCCGGGCACTCATGCTCGCGCTG 549
Qy 59 GlnGluAspLysValLeuIleThrGlnLeuIleGlnHisSerAla---LysHisLysVal 77
Db 550 CAGAAGGGTAACTCGCGCGGTGCGCGCTCTTCAAGCGCGGAGACCGGACCCGCTTC 609
Qy 78 AspIleAspCys---SerGlnLysValValValTyrAspGlnSerGlnAsp---Val 95
Db 610 ACCCGGCGCTGGCACCCGACACAGTGTGCTCTTACGACGAGACGACGACTGGAAC 669
Qy 96 AlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPhe 115
Db 670 GAGAAATACGGGCGGCGGAGTGTGCTCGGCTCTCTCAAGAAAGCTCAAGACAGAGGGC 729
Qy 116 AsnSerValHisLeuLeuAlaGlyIlePheAlaGlnPheSerArgCysPheProGlyLeu 135
Db 730 TGCAGGCGCTTACTGAGAAAGTGTGCTTCAAGATTCACAAAGCAGATCTCCCTGCAT 789
Qy 136 CysGlnGlyLysSerThrLeuValProThrCys---IleSerGlnProCysLeuProVal 154
Db 790 TGGGAG---ACCATCTAAGCGGCTGTGTAGACGACGCTCGCGCGCTTGCAGTG 843
Qy 155 AlaAsnIleGly----- 158
Db 844 CTGGGGCTGGGGGCTGGGATCAGCTGATCTTCTCGGACATCAGTGTGACTT 903
Qy 158 ----- 158
Db 904 GACGAGACCCCAATAGTGCAACAGACTCGATGATGCTCCGCTGCCAACAGCAGACT 963
Qy 159 ---ProThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsn 176
Db 964 TCCTTCCAGTGAAGATCTTCCCTTCTTACTTGGGCTGTGCAAGATCCACCAAC 1023
```

QY 177 LysGluLeuIleGlnGlnAenglyIleGlyTyrValLeuAsnAlaSerTyrrhCysPro 196
DB 1024 TTGGAGCTGTGGAGAAATTCGGCATCAAGTACATCTTGAACCGTCAACCCCAATTTGCGG 1083
QY 197 LysPro---AspPheIleProGlnSerHisPheLeuArgValProValAsnAspSerPhe 215
DB 1084 AATCTCTTGAGAACCGACGAGAGTTTAATACAGCAATCCCATCTCGATCCTGG 1143
QY 216 CysGluLysIleLeuProTribLeuAspLysSerValAspPheIleGluValAlaVala 235
DB 1144 AGCCAAACCTGTCACGATTTTCCCTGAGGCGCATTTCTTCAATAGTGAAGCCCGGGG 1203
QY 236 SerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAla 255
DB 1204 AAGAACTGTGTCTTCTTGTACATTCCTGGCTGGCAATGACCGCTCAGCTGAGCT 1263
QY 256 IleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLys 275
DB 1264 GTGGCTTACCTTATTCGAAAGCTCAATCTGTGCATGACGATCCCTATGCAATTCGAA 1323
QY 276 GlnLysArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuAspTyrGlu 295
DB 1324 ATGAAATAATCCAAACATATCCCTAACTTCACTTCAATGAGTGTGAGCTGCACTTGAG 1383
QY 296 LysLysIle 298
DB 1384 AGGACGCTG 1392
RESULT 6
US-09-023-655-946
Sequence 946, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Sellhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 946:
SEQUENCE CHARACTERISTICS:
LENGTH: 2109 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G1418933
US-09-023-655-946
Alignment Scores:
Prod. No.: 2, 91e-51
Score: 467.50
Percent Similarity: 53.568
Best Local Similarity: 34.064
Query Match: 29.894
Gaps: 4
Length: 2109
Matches: 110
Conservative: 63
Mismatch: 105
Indels: 45
Gaps: 7
US-10-029-345A-109_COPY_1_302 (1-302) x US-09-023-655-946 (1-2109)
QY 19 LeuGlnSerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsn 38
DB 430 CTGAGACTGGGCAACGAGGCGCTGCTGATGAGATCGCGCGGCGGAGCTATACGAG 489
QY 39 ThrSerHisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeu 58
DB 490 TCGTCGACATCAGTGGCCATCAAGTGGCCATCCCGGCGCATGCTGCGGCGCTG 549
QY 59 GlnGlnAspLysValLeuIleThrGluLeuIleGlnHisSerAla---LysHisLysVal 77
DB 550 CAGAGGATTAACCTGCGCGGCGCGCTCTTACGCGGCGGAGAGACCGGACCGCTTC 609
QY 78 AspIleAspCys---SerGlnLysValValTyrAspGlnSerSerGlnAsp---Val 95
DB 610 ACCCGCGCTGTGGACACGACAGAGTGTCTTACGACGAGAGAGAGAGAGAGAGAG 669
QY 96 AlaSerLysSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPhe 115
DB 670 GAGAAATCGGCGCGCGCGCTGCTGCGGCGCTGCTCAAGAGCTCAAGAGAGAGAG 729
QY 116 AsnSerValHisLeuLeuValGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeu 135
DB 730 TCGCGGCGCTTCTTACTGGAAGTGGCTTCAAGTAATTCAGACCGAGTCTCCGCAAT 789
QY 136 CysGlnGlyLysSerThrLeuValProThrCys---IleSerGlnProCysLeuProVal 154
DB 790 TGGGAG---ACCAATCTAAGACGCGCTGTGAGAGAGAGCTGCGCGCGCTGCGAGT 843
QY 155 AlaAsnIleGly----- 158
DB 844 CTGGGCTGCGGCGCTGCGGATCAGCTGACTCTTCTCGGACATGAGTGAAGCTT 903
QY 158 ----- 158
DB 904 GACCGAGACCCCAATGTCACACAGCTCGGATGATGTCCTGTCACACGACGCT 963
QY 159 ---ProThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgPheValLeuAsn 176
DB 964 TCTTCTCCAGTGGAGATCTTGGCTTCTTACTTGGCTGTGCCAAAGCTCCACCAAC 1023
QY 177 LysGluLeuIleGlnGlnAenglyIleGlyTyrValLeuAsnAlaSerTyrrhCysPro 196
DB 1024 TTGGAGCTGTGGAGAAATTCGGCATCAAGTACATCTTGAACCGTCAACCCCAATTTGCGG 1083
QY 197 LysPro---AspPheIleProGlnSerHisPheLeuArgValProValAsnAspSerPhe 215
DB 1084 AATCTCTTGAGAACCGACGAGAGTTTAATACAGCAATCCCATCTCGATCCTGG 1143
QY 216 CysGluLysIleLeuProTribLeuAspLysSerValAspPheIleGluValAlaVala 235
DB 1144 AGCCAAACCTGTCACGATTTTCCCTGAGGCGCATTTCTTCAATAGTGAAGCCCGGGG 1203
QY 236 SerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAla 255
DB 1204 AAGAACTGTGTCTTGTGTACATTCCTGGCTGGCAATGACCGCTCAGCTGAGCT 1263
QY 256 IleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLys 275

Db 291 GCGGGCTACATCTTGAGTTGGGTCAACGCGCCTGTAAACCATCTGTGGCGCGG--- 347
 Oy 59 GInGInAePlyValLeuLeuTherGluLeuLeu-----GlnHisSerAlaIysHis 76
 Db 348 GGTAAAGGGCTCGGTAGAGCTGAGACAGATCTCGCCCGGAGAGAGAGTATCGCGCCG 407
 Oy 77 ValAaPpIleAspCysSerGlnIysValValValIYrAspGlnSerSerGlnAspValAla 96
 Db 408 TTGGCCTCGGCCTTACTCTGAGCGGTCTACATCGTTCAGTCAAGACGAGGAGCGCCGCGAG 467
 Oy 97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGlnIySerPhe--- 115
 Db 468 ACCCTTCGGAGAGACAGACCGCTGTCTGGTGGTGGAGAGCGCCTGCGCCGACCGCGAG 527
 Oy 116 --AasSerValHisLeuLeuValaGlyLysPheAlaGluPheSerArgCysPheProGly 134
 Db 528 CGCAGCCAGATCTGCTGCTCAAGAGCGGCTATGAGAGAGTTTCTCCGAGTACACAGAA 587
 Oy 135 LeuGlyGlnGlyLysSerThreLeu-----ValProThrCysIleSer 148
 Db 568 TTCTGTCTTAAACCAAGAGCCTGGCAGCATCCACCCCGGTTTCCCGACGCGCACAA 647
 Oy 149 GlnPro-----Cys--LeuProValAlaAsnIle--GlyPro 159
 Db 648 GAGCCTTGGACCTGGGCTGACAGCTCTGTGGAGCCCACTACACAGACAGAGGGGTCTT 707
 Oy 160 ThrArgIleLeuProAsnLeuIYrLeuGlyCysGlnIArgAspValLeuAsnIySerLeu 179
 Db 708 GTGAGATCTCTTCCCTTCTTCTTCACTCGGACAGTGCCTACATGCTGCCCGAGAGACATG 767
 Oy 180 IleGlnGlnAsnGlyIleGlyLysValLeuAsnAlaSerIYrThrCysProIyAsp 199
 Db 768 CTGACAGCGCTCGGAGATACAGGCTCTGTGAATGTCTCTGGATGTCGCCAAC--CAC 824
 Oy 200 PheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlnIySile 219
 Db 825 TTGTGAAGACACTATCACTAGTCAAGTGCATCCAGTGAAGATAACCAAGGCCAGCATCT 884
 Oy 220 LeuProTyrLeuAspLysSerValAspPheIleGlySalAlaLysAlaSerAsnIyCys 239
 Db 885 AGCTCTGTTCATGGAAGCCATAGATAGTACATGCGCGTGAAGAGATGCGCGTGGCGC 944
 Oy 240 ValLeuValHisCysLeuValGlyLysSerArgSerAlaThrIleAlaIleAlaIYrIle 259
 Db 945 GTGCTGTGTGCACTGCCAGGCGGAGATCTCGCGTGGCCACATCTGCTGGCTCACTG 1004
 Oy 260 MetLysArgMetAspMetSerLeuAspGlnAlaIYrArgPheValIyGlnIyAspArgPro 279
 Db 1005 ATGATGAAGAAACGGGTGAGGCTGAGAGAGCCTTCGATTCGTTAAGCACAGCGCGCAGC 1064
 Oy 280 ThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIYrGlnIyLysIle 298
 Db 1065 ATCATCTCGCCCACTTCAGCTTATGGAAGACGTGCTGCGAGTTCAGTCCAGGTG 1121

RESULT 8
 US-09-371-671B-10
 ; Sequence 10, Application US/09371671B
 ; Patent No. 6548743
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheen, Jen
 ; APPLICANT: Chiu, Man-Ling
 ; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING A
 ; TITLE OF INVENTION: DUAL-SPECIFICITY MAPK PHOSPHATASE AND USES THEREOF
 ; FILE REFERENCE: 00786/370002
 ; CURRENT APPLICATION NUMBER: US/09/371,671B
 ; CURRENT FILING DATE: 1999-08-10
 ; PRIOR APPLICATION NUMBER: 60/155,934
 ; PRIOR FILING DATE: 1998-01-14
 ; PRIOR APPLICATION NUMBER: 60/095,938
 ; PRIOR FILING DATE: 1998-08-10
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 4.0


```

; SEQ ID NO 10
; LENGTH: 944
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-371-671B-10

Alignment Scores:
Pred. No.:      3,18e-46      Length:      944
Score:          425.00       Matches:     104
Percent Similarity: 50.17%   Conservative: 48
Best Local Similarity: 34.32% Mismatches:    129
Query Match:     27.17%     Indels:       22
DB:              4           Gaps:         5

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-371-671B-10 (1-944)

QY      15 LeuValAlaLeuIleuGluSerGlyThrGluValLeuIleuLeuLeuAspSerArgProPhe 34
Db      39 CTGGCGGCGCTGTGGAGAGAGCGCGCCGACATGCCTGTGTGGATTGTGCTCCTTC 98
QY      35 ValGluTyraAnthrSerHisIleLeuGluAlaIleAsnIleAsnCysSerLysMet 54
Db      99 TTGCTTTCAACGCCGACCACATCGGGGCTCATGAACCTGCGCTTCAAGCACATCGTG 158
QY      55 LysArgArgLeuGlnGlnAspLysValLeuIleThrGluLeuIleGlnHisSerAlaLys 74
Db      159 CGCGCGCGCGCCCAAGGCGCCATGGCGCTGAGCATTCCTGCCCAACGCTAACTGCTC 218
QY      75 HisLysValAspIleAspCysSerGlnLysValValValTyraAspGlnSerSerGlnAsp 94
Db      219 GGCGCGCTGTGGCCGGAGGCTTAACACGCGTGTGTGCTGTGAACAAGCGCAGCGCTCC 278
QY      95 ValAlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGluLysLysLeu-----Glu 112
Db      279 CTGACGCGGCGCCAAAGCGCGACGCGACCTGCGCCCTGGCCGCGGCGCGCTCTGCCAGAG 338
QY      113 LysSerPheAsnSerValHisLeuLeuAlaGlyLysPheAlaGluPheSerArgCysPhe 132
Db      339 GCAGCGCTCCACTCAAGACTCTTCTTCTCCAAAGAGATATAGAAGCTTTTCGGCTTCTGC 398
QY      133 ProGlyLeuGlyGluGlyLysSerThrLeuValProthrCysIleSerGlnPro----- 150
Db      399 CCTGAGCTGTGCAGCAACACAGTCCACC-----CCACGAGGAGCTCAACCTCCCCTGAGT 452
QY      151 -----CysLeuProValAlaAsnIle----- 157
Db      453 ACTAGTGTGCTGACAGTGCAGATCCGAGATCCGATGACGCTCGTAGTACCCCTCTACGAT 512
QY      158 -----GlyProThrArgIleLeuProAsnLeuTyrlGluLysGlnIndArgAspValLeu 175
Db      513 CAGGAGGAGTGTGTCAGTGAAGATCTGTCTCTCTGACTGAGGACAGTGCCTTCAACGCTTCT 572
QY      176 AsnLysGluLeuIleGlnGlnAsnGlyLleGlyTyraValLeuAsnAlaSerTyrrThrCys 195
Db      573 CGGAAGGATATGCTGACGCGCTTGGGATACCGCCTTGATCAACAGCTCTACAGCAATTGT 632
QY      196 ProLysProAspPheIleProGlnSerHisPheLeuArgValProValAsnAspSerPhe 215
Db      633 CTTRAAC---CACTTTGAGGCTCACTACCAATPACAAGACATCCCTGTGAGAGACAACAC 689
QY      216 CysGluLysIleLeuProThrPleuAspLysSerValAspPheIleGlyValAlaLysAla 235
Db      690 AAGGACAGACATAGCTCTGCTGTTCAAAGAGGATTAATGACTTCAATAGACATCAAGAT 749
QY      236 SerAsnGlyCysValLeuValHisCysLeuAlaGlyLysSerArgSerAlaThrIleAla 255
Db      750 GCTGAGAGGAGAGTGTGTTCACTTGCACAGCGCGCATCTCCCGGTGAGCCACCATCTGC 809
QY      256 IleAlaTyrlleMetLysArgMetAspMetSerLeuAspGlnAlaTyraCspPheValLys 275
Db      810 CTTGCTTACTCAATGAGGACTTAACCGGGTAAAGCTGACAGAGGCTTTGAGATTGTGACAG 865
QY      276 GluLysArgProThrIleSerProAsnPheAsnPhenLeuGluLysLeuLeuAspArgGlu 295

```

```

Db      870  CAGAGCGGAGTATCATCTCCCGGAATTGAGTTGATGGCCAGCTGCTGACATTGAG  929
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Oy      296  LysLysIle 298
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      930  TCCCAAGTG 938
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

RESULT 9
US-08-530-290-11
/ Sequence 11, Application US/08530290
/ Patent No. 5958721
/ GENERAL INFORMATION:
/ APPLICANT: Marshall, Christopher John
/ APPLICANT: Ashworth, Alan
/ APPLICANT: Hughes, David Anthony
/ TITLE OF INVENTION: Methods for Screening of Substances for
/ TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
/ NUMBER OF SEQUENCES: 24
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/530,290
/ FILING DATE: 14-DEC-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO PCT/GB94/00694
/ FILING DATE: 31-MAR-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9402573.1
/ FILING DATE: 10-FEB-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9307250.2
/ FILING DATE: 07-APR-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 084611-000000US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1238 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ US-08-530-290-11

Alignment Scores:
Pred. No.: 5,02e-46 Length: 1238
Score: 425.00 Matches: 99
Percent Similarity: 53.85% Conservative: 62
Best Local Similarity: 33.11% Mismatches: 116
Query Match: 27.17% Indels: 22
DB: 2 Gaps: 8

US-10-029-345A-109_COPY_1_302 (1-302) x US-08-530-290-11 (1-1238)
Oy      19  leuGlSerGlThrGlLysValIleuLeuIleAspSerArgProPheValGIuTyrAsn 38
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      168  CTCGCCAGCGGC---GCGAAGTGCCTGCTGCTGACATCGACACCGCTTCTCGCGCACAAGC  224

```

```

Qy 39 ThrsrhisileuagualaleasnleasnCyserlysluMetlysrArgyleu 58
Db 225 GCGGGCTACATCTCGGTTCAGCGCTGTAACGCTGTAACGCTGCGCGGCGG--- 281
Qy 59 GlnGlnAspLysValLeuileThrgluLeuile-----GlnHisSerAlaLysHisLys 76
Db 282 GCTAAGAGGCTCCGTGAGCTGAGACATCTGCGCGCGGAGAGAGGTACGCGCGCGC 341
Qy 77 ValAspIleAspCyserGlnLysValValValTyrrAspGlnSerSerGlnAspValAla 96
Db 342 TTGGCTCCCGGCTCTACTCGCGCGGTATCGCTTACAGACGACCGCGCGCGCGCGAG 401
Qy 97 SerLeuSerSerAspCyserPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPhe--- 115
Db 402 AGCTCCCGAGAGACAGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 461
Qy 116 ---AsnSerValHisLeuLeuAlaGlyValPheAlaGluPheSerArgCysPheProGly 134
Db 462 CGCACCGACATCTGCTGCTCAAGCGGCTATGAGAGGTTTCTCCGAGTACCCAGAA 521
Qy 135 LeuCyserGlnLysSerThrLeu-----ValProThrCysIleSer 148
Db 522 TTCTGTTCTAAACCAAGCGCTGCGAGCATCCACCGCGGTTCGCCCGAGCGCAC 581
Qy 149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db 582 GAGCCCTTGAGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641
Qy 160 ThrArgIleLeuProAsnLeuTyrrLeuGlyCysGlnArgAspValLeuAsnLysGlnLeu 179
Db 642 GTGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 701
Qy 180 IleGlnGlnAsnGlyIleGlyTyrrValLeuAsnLysSerTyrrCysProLysProAsp 199
Db 702 CTGAGACCGCTGCGGATACCGCTGCTGTAATGCTCTCTCTGCGACTGCCAAC---CAC 758
Qy 200 PheIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIle 219
Db 759 TTTCAGAGACACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 818
Qy 220 LeuProTrpLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCys 239
Db 819 AGCTCCGTGTTCAATGAGAGCATGAGTACATGATGATGATGATGATGATGATGATG 878
Qy 240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrrIle 259
Db 879 GTGCTGTGCTACCTCGACGCGGCGATCTCGCGGTGCGGCGGCGGCGGCGGCGGCGG 938
Qy 260 MetLysArgMetAspMetSerLeuAspGluAlaTyrrArgPheValLysGlnLysArgPro 279
Db 939 ATGATGAAGAAACGAGTGAAGCTGAGAGGCTTCGAGTTGTTAAGCAGCGCGCGAGC 998
Qy 280 ThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrrGlnLysLysIle 298
Db 999 ATCATCTCGCCCACTTCACTCATGAGGCGAGCTGCGAGTTCGAGTCCAGGTG 1055

RESULT 10
US-09-702-705-803
; Sequence 803: Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedrick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLES OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14

```

```

; CURRENT APPLICATION NUMBER: US/09/702, 705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 803
; LENGTH: 1238
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-803

Alignment Scores:
Pred. No.: 5.02e-46 Length: 1238
Score: 425.00 Matches: 99
Percent Similarity: 53.85% Conservative: 62
Best Local Similarity: 33.11% Mismatches: 116
Query Match: 27.17% Indels: 22
DB: Gaps: 8

US-10-029-345a-109_copy_1_302 (1-302) x US-09-702-705-803 (1-1238)
Qy 19 LeuGlnSerGlyThrGlnLysValLeuLeuIleAspSerArgPheValGlnTyrrAsn 38
Db 168 CTGCGAGCGGAC---GCGATGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 224
Qy 39 ThrsrhisileuagualaleasnleasnCyserlysluMetlysrArgyleu 58
Db 225 GCGGGCTACATCTCGGTTCAGCGCTGTAACGCTGTAACGCTGCGCGGCGG--- 281
Qy 59 GlnGlnAspLysValLeuileThrgluLeuile-----GlnHisSerAlaLysHisLys 76
Db 282 GCTAAGAGGCTCCGTGAGCTGAGACATCTGCGCGGAGAGAGGTACGCGCGCGC 341
Qy 77 ValAspIleAspCyserGlnLysValValValTyrrAspGlnSerSerGlnAspValAla 96
Db 342 TTGGCTCCCGGCTCTACTCGCGCGGTATCGCTTACAGACGACCGCGCGCGCGAG 401
Qy 97 SerLeuSerSerAspCyserPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPhe--- 115
Db 402 AGCTCCCGAGAGACAGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
Qy 116 ---AsnSerValHisLeuLeuAlaGlyValPheAlaGluPheSerArgCysPheProGly 134
Db 462 CGCACCGACATCTGCTGCTCAAGCGGCTATGAGAGGTTTCTCCGAGTACCCAGAA 521
Qy 135 LeuCyserGlnLysSerThrLeu-----ValProThrCysIleSer 148
Db 522 TTCTGTTCTAAACCAAGCGCTGCGAGCATCCACCGCGGTTCGCCCGAGCGCAC 581
Qy 149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db 582 GAGCCCTTGAGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641
Qy 160 ThrArgIleLeuProAsnLeuTyrrLeuGlyCysGlnArgAspValLeuAsnLysGlnLeu 179
Db 642 GTGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 701
Qy 180 IleGlnGlnAsnGlyIleGlyTyrrValLeuAsnLysSerTyrrCysProLysProAsp 199
Db 702 CTGAGACCGCTGCGGATACCGCTGCTGTAATGCTCTCTCTGCGACTGCCAAC---CAC 758
Qy 200 PheIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIle 219
Db 759 TTTCAGAGACACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 818
Qy 220 LeuProTrpLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCys 239
Db 819 AGCTCCGTGTTCAATGAGAGCATGAGTACATGATGATGATGATGATGATGATGATG 878
Qy 240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrrIle 259
Db 879 GTGCTGTGCTACCTCGACGCGGCGATCTCGCGGTGCGGCGGCGGCGGCGGCGGCGG 938
Qy 260 MetLysArgMetAspMetSerLeuAspGluAlaTyrrArgPheValLysGlnLysArgPro 279

```

```
Db 939 ATGATGAGAAACGGGAGAGCGGAGGAGCGCTTGAGTTGCTTAAGACGCGCCAGC 998
Qy 280 ThrilleserProAnpheaPheLeuGlyGlnLeuLeuAspTyrGlnLysIle 298
Db 999 ATCATCTCGCCCAACTTCACCTTCATGCGGACGCTGCTGACGTTCCAGTCCAGGTG 1055

RESULT 11
US-09-736-457-803
/ Sequence 803, Application US/09736457
/ Patent No. 6509448
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ APPLICANT: Wang, Aijun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C15
/ CURRENT APPLICATION NUMBER: US/09/736.457
/ CURRENT FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 1864
/ SOFTWARE: FaastSeq for Windows Version 3.0
/ SEQ ID NO 803
/ LENGTH: 1238
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-736-457-803

Alignment Scores:
Pred. No.: 5,02e-46 Length: 1238
Score: 425.00 Matches: 99
Percent Similarity: 53.85% Conservative: 62
Best Local Similarity: 33.11% Mismatches: 116
Query Match: 27.17% Indels: 22
Gaps: 8
DB: 4

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-736-457-803 (1-1238)
Qy 19 LeuGlnSerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnTyrAsn 38
Db 168 CTGCCGAGCGGC---GGCAAGTGCCTGCTGCTGAGCTGCAGACCGTTCTCGGCGCACAGC 224
Qy 39 ThrSerHisIleLeuGlnLysIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 58
Db 225 GCGGCGCTACATCTTACGTTGCTGTCACAGTGCCTGTAAACCATTCGCGCGCGCGG--- 281
Qy 59 GlnGlnAspLysValLeuIleThrGlnLysIle-----GlnHisSerAlaLysHisLys 76
Db 282 GCTAAGGCTCCTGAGCTGAGCAGATCTGCGCCCGCCAGAGAGAGATACCGCCCGC 341
Qy 77 ValAspIleAspCysSerGlnLysValValTyrAsnIleAsnIleAsnIleAsnIleAsnIleAsn 96
Db 342 TTGCGCTCCGCGCTCTACTGCGCGCTCATGCTGTACAGACGAGCGGCGCGCGCGAG 401
Qy 97 SerLeuSerSerArgPheLeuThrValLeuLeuGlnLysIleLeuGlnLysSerPhe--- 115
Db 402 AGCTCTCGCAGAGACACCGCTGCTGCTGCTGAGCGGCTGCGCCGCAACGCCGAG 461
Qy 116 ---AsnSerValHisLeuLeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGly 134
Db 462 CGCAGCGACATCTGCGCTGCTCAAGGCGGCTATGAGAGGTTTCTCCGAGTACCCAGAA 521
Qy 135 LeuGlnGlnLysSerThrLeu-----ValProThrCysIleSer 148
Db 522 TTCTGTTCTAAACCAAGGCGCTGCGAGCATCCACCCCGGTTCCCGCCAGCGCCACA 581
```

```
Qy 149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db 582 GAGCCCTTGAGACTGACCTGACGCTCTGTGGAGCCCACTACAGACCAAGAGGCTCT 641
Qy 160 ThrArgIleLeuProAnpLeuTyrIleuGlyCysGlnArgAspValLeuAnLysGlnLeu 179
Db 642 GTGAGATCTCTTCCCTTCCCTTACCTCGGACAGTCCCTACCATGCTCCCGGAGAGCATG 701
Qy 180 IleGlnGlnAsnGlyIleGlyTyrValLeuAnpAlaSerTyrThrAspProLysProAnp 199
Db 702 CTGAGCGCCCTGGGACACCGCTCTGTGAATGCTCTCCGACGCTGCCAAC---CAC 758
Qy 200 PheIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIle 219
Db 759 TTGGAAGACACTATAGTACAAAGTCCATCCAGTGAAGATACCAACAAAGCCGACATC 818
Qy 220 LeuProTyrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCys 239
Db 819 AGCTCTGCTCATGAGACCATAGATCATGATGATGATGATGATGATGATGATGATGATGATG 878
Qy 240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 259
Db 879 GTGCTGTGCTCAGTGCAGGCGGACATCTGCGGTCGCGCCACCATCTGCTGCGCTGCTG 938
Qy 260 MetLysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlnLysArgPro 279
Db 939 ATGATGAGAAACGGGTGAGGCTGAGAGGCTTGAGATTCCTTAAGACGCCGCCAGC 998
Qy 280 ThrilleserProAnpheaPheLeuGlyGlnLeuLeuAspTyrGlnLysIle 298
Db 999 ATCATCTCGCCCAACTTCACCTTCATGCGGACGCTGCTGACGTTCCAGTCCAGGTG 1055

RESULT 12
US-09-614-124B-803
/ Sequence 803, Application US/09614124B
/ Patent No. 6630574
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ FILE REFERENCE: 210121.478C9
/ CURRENT APPLICATION NUMBER: US/09/614.124B
/ CURRENT FILING DATE: 2001-07-11
/ NUMBER OF SEQ ID NOS: 1668
/ SOFTWARE: FaastSeq for Windows Version 3.0
/ SEQ ID NO 803
/ LENGTH: 1238
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-614-124B-803

Alignment Scores:
Pred. No.: 5,02e-46 Length: 1238
Score: 425.00 Matches: 99
Percent Similarity: 53.85% Conservative: 62
Best Local Similarity: 33.11% Mismatches: 116
Query Match: 27.17% Indels: 22
Gaps: 8
DB: 4

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-614-124B-803 (1-1238)
Qy 19 LeuGlnSerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnTyrAsn 38
Db 168 CTGCCGAGCGGC---GGCAAGTGCCTGCTGCTGAGCTGCAGACCGTTCTCCGCGCACAGC 224
Qy 39 ThrSerHisIleLeuGlnLysIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 58
```


Db 939 ATGATGAAGAAACGGGTGAGGCTGAGAGCCCTTCAGTTCCGTTAAGACAGCCGCGAC 998
Qy 280 ThrIleSerProAnpPheAnpPheLeuGlyValLeuLeuAspTyrGlyValLeu 298
Db 999 ATCATCTGCGCCAACTTCACTTCACTGAGGCGAGCTGCGAGTTCGAGTCCAGAGT 1055

RESULT 14
US-09-589-184-803
Sequence 803, Application US/09589184
Patent No. 6686447
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaltanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C8
CURRENT APPLICATION NUMBER: US/09/589,184
CURRENT FILING DATE: 2000-06-05
NUMBER OF SEQ ID NOS: 827
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 803
LENGTH: 1238
TYPE: DNA
ORGANISM: Homo sapiens
US-09-589-184-803

Alignment Scores:
Pred. No.: 5,02e-46 Length: 1238
Score: 425.00 Matches: 99
Percent Similarity: 53.85% Conservative: 62
Best Local Similarity: 33.11% Mismatches: 116
Query Match: 27.17% Indels: 22
Gaps: 8

US-10-029-345a-109_copy_1_302 (1-302) x US-09-589-184-803 (1-1238)

Qy 19 LeuGluSerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 38
Db 168 CTCGCGAGCGCGC---GCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 224
Qy 39 ThrSerHisIleLeuGluAlaIleAsnIleAsnCySerIleuMetIleuMetIleuMet 58
Db 225 GCGGGCTACATCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 281
Qy 59 GluGlnAspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 76
Db 282 GCTAAGGCTCTCGAGCTCGAGCTCGAGCTCGAGCTCGAGCTCGAGCTCGAGCTCGAGCT 341
Qy 77 ValAspIleAspCySerGlnValValValValValValValValValValValValVal 96
Db 342 TTGCGCTCCGCGCT 401
Qy 97 SerLeuSerSerAspCyPheLeuThrValLeuLeuGlyIleuGluIleuGluIleuGlu 115
Db 402 AGCTCTCGAGGAGAGACACCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 461
Qy 116 ---AsnSerValHisLeuLeuAlaGlyIleuPheAlaGluPheSerArgCyPheProGly 134
Db 462 CGACACCGACATCTGCTCTCAAGGCGCTAAGAGGTTTCTCTCGAGTCCAGACCCAGAA 521
Qy 135 LeuCySerGluIleuValSerThrLeu-----ValProThrCyIleSer 148
Db 522 TTCTGTCTTAAACCAAGCGCTCTGAGCAGCATCCACCCCGGTTCCCGCCAGCGCAC 581
Qy 149 GlnPro-----CyLeuProValAla-----AsnIleGlyPro 159
Db 582 GAGCCCTTGGACCTGAGCTGAGCTCTGAGGAGCCCACTACACGACGAGAGGCTCT 641

Qy 160 ThrArgIleLeuProAnpLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 179
Db 642 GTGAGATCT 701
Qy 180 IleGlnGlnAnGlyIleGlyIleValLeuAsnAlaSerTyrThrCyProIleProAsp 199
Db 702 CTGAGCGCTCTGAGCAGCGCTCTGTAATGTCTCTCTGAGCTCCCAAC---CAC 758
Qy 200 PheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCyGluValIle 219
Db 759 TTGAAGACACTATCACTCAAGTGCATCCAGTGAATATTAACCAAGCCGACATC 818
Qy 220 LeuProThrLeuAspIleValSerValAspPheIleGluValAlaValAsnAlaSer 239
Db 819 AGCTCTGTCTTACATGAGCATATGATGATGATGATGATGATGATGATGATGATGATG 878
Qy 240 ValLeuValHisCySerValGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 259
Db 879 GTGCTGTGACATGCGAGCGGCGATCTCGCGTCCGACCATCTCTGCTCTCTCTCTCTCT 938
Qy 260 MetLeuArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluValArgPro 279
Db 939 ATGATGAAGAAACGGGTGAGGCTGAGAGGCTTCGAGTTCCGTTAAGACAGCCGCGAC 998

Qy 280 ThrIleSerProAnpPheAnpPheLeuGlyValLeuLeuAspTyrGlyValLeu 298
Db 999 ATCATCTGCGCCAACTTCACTTCACTGAGGCGAGCTGCGAGTTCGAGTCCAGAGT 1055

RESULT 15
US-09-702-705-825
Sequence 825, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaltanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 825
LENGTH: 2064
TYPE: DNA
ORGANISM: Homo sapiens
US-09-702-705-825

Alignment Scores:
Pred. No.: 1.18e-45 Length: 2064
Score: 425.00 Matches: 100
Percent Similarity: 54.85% Conservative: 64
Best Local Similarity: 33.44% Mismatches: 113
Query Match: 27.17% Indels: 22
Gaps: 9

US-10-029-345a-109_copy_1_302 (1-302) x US-09-702-705-825 (1-2064)

Qy 19 LeuGluSerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 38
Db 511 CTCGCGAGCGCGC---GCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 567
Qy 39 ThrSerHisIleLeuGluAlaIleAsnIleAsnCySerIleuMetIleuMetIleuMet 58
Db 568 GCGGGCTACATCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 624

